

From: Mertz, Prema
Sent: Saturday, October 19, 2002 2:53 PM
To: STIC-Biotech/ChemLib
Subject: 09/851,595

Please search SEQ ID NO:11 with protein databases.

Thanks,

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

10/23/02
ABH
1-AA

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52 ; Search time 37 Seconds
(without alignments)

2902.929 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068

Sequence: 1 MPSPGRLALMLCALCASR.....GGLSGGGGFQPSGLAFASHV 967

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4067	80.2	847	AAV53575	Human gonadotropin
2	2886.5	57.0	691	AAV53575	Novel human diapo
3	2525.5	49.8	907	AAW93889	Human Hg38 protein
4	2525.5	49.8	907	AAV90682	Human G protein-co
5	2519.5	49.7	907	AAV90687	Human mutant G pro
6	2450	48.3	497	AAV53574	Human gonadotropin
7	2035.5	40.1	951	AAV53574	Human RBCAR polype
8	2033	40.1	395	AAV53574	A human guanosine
9	2030.5	40.1	949	AAV53574	Human AOMF05 prote
10	2030.5	40.1	949	AAV53574	Human AOMF05 prote
11	2030.5	40.1	951	AAV53574	Human AOMF05 prote

12	2030.5	40.1	951	AAV53574	Human AOMF05 prote
13	2030.5	40.1	951	AAV53574	Novel human secret
14	2027.5	39.4	951	AAV53574	Human LGR4 protein
15	1996	37.5	693	AAV53574	Human Hg38 protein
16	1901	37.5	693	AAV53574	Human LGR5 protein
17	1367	27.0	266	AAV53574	Human ORF1347
18	1206	23.8	230	AAV53574	Human protein sequ
19	1140	22.5	644	AAV53574	Human 7-transmembr
20	1102	21.7	631	AAV53574	Portion of Tingo-7
21	1010.5	19.9	1300	AAV53574	Drosophila melanog
22	901	17.8	171	AAV53574	Human gonadotropin
23	688.5	13.6	861	AAV53574	D. melanogaster pe
24	536	10.4	141	AAV53574	Human Leu-rich rep
25	528.5	10.4	699	AAV53574	N-terminal of LH r
26	525	10.4	120	AAV53574	Human expressed po
27	525	10.4	120	AAV53574	Human secreted pro
28	523	10.3	695	AAV53574	FSHR. Homo sapien
29	519.5	10.3	695	AAV53574	Rat testicular lut
30	518	10.2	689	AAV53574	FSH receptor. Hom
31	517.5	10.2	695	AAV53574	N-terminal of LH r
32	517	10.2	695	AAV53574	N-terminal of LH r
33	517	10.2	695	AAV53574	N-terminal of LH r
34	515.5	10.2	695	AAV53574	N-terminal of LH r
35	512	10.1	700	AAV53574	N-terminal of LH r
36	511	10.1	692	AAV53574	N-terminal of LH r
37	510	10.1	692	AAV53574	N-terminal of LH r
38	508.5	10.0	695	AAV53574	N-terminal of LH r
39	507	10.0	698	AAV53574	N-terminal of LH r
40	502.5	9.9	202	AAV53574	Human gonadotropin
41	497	9.8	696	AAV53574	N-terminal of LH r
42	492.5	9.7	764	AAV53574	N-terminal of LH r
43	491.5	9.7	829	AAV53574	Canine thyroid stl
44	489.5	9.7	696	AAV53574	Drosophila melanog
45	486.5	9.6	696	AAV53574	N-terminal of LH r

ALIGNMENTS

RESULT 1
AAV53575
ID AAV53575 standard; Protein: 847 AA.
AC AAV53575;
DT 15-FEB-2000 (first entry)
XX
DE Human gonadotropin receptor partial sequence #5.
KW Human: gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
XX
OS Homo sapiens.
XX
PN EP950711-A2.
PD 20-OCT-1999.
PF 02-FEB-1999; 99EP-0200303.
PR 06-FEB-1998; 98EP-0200357.
PR 27-JUL-1998; 98EP-020519.
PR 24-SEP-1998; 98EP-0205213.
PA (ALKU) AKZO NOBEL NV.
PI Van Der Spek PI, Heikoop JC;
XX WPI: 1999-563673/48.
DR N-FSD: AAZ40461.
XX New 7 transmembrane gonadotropin receptors, useful for screening for

PT hormone analogs and drugs -

PS Claim 1: Page 24-28; 38pp; English.

CC Sequences AAY53571-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.

SQ Sequence 847 AA:

Query Match 80.2%; Score 4067; DB 20; Length 847;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 2; Indels 96; Gaps 1;

QY 72 DLSMNNLELPDGLFHHHRLFEELRLSGNHLSHI PGQAFSGSLYSLKTIIMQNNQIGIPA 131
DB 29 DLSMNNLELPDGLFHHHRLFEELRLSGNHLSHI PGQAFSGSLYSLKTIIMQNNQIGIPA 88
QY 132 EALMELPSLOSRLDANLISLVERSEGLSLRHLMLDNLTEIPVALNMLPALQAM 191
DB 89 EALMELPSLOS-----
QY 192 TLALNRSHTPDYAFONLISLVYLHNNRIQHLGTHSEGLHNLETLNLNKKLOEFPV 251
DB 100 -----
QY 252 AIRTLGRLOELGFHNNNIKAIPKAFMGNPILOTIFYNPIQFGRSAFOYLPKLTLS 311
DB 113 AIRTLGRLOELGFHNNNIKAIPKAFMGNPILOTIFYNPIQFGRSAFOYLPKLTLS 172
QY 312 LMGANDIOEPDLKGTSTLEITLTPRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLH 371
DB 173 LMGANDIOEPDLKGTSTLEITLTPRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLH 232
QY 372 RQKLEELIGLQHNRIWEIGADTFSSLOALDLSMNAIRSHHPAFSTLSLVKLDLTD 431
DB 233 RQKLEELIGLQHNRIWEIGADTFSSLOALDLSMNAIRSHHPAFSTLSLVKLDLTD 292
QY 432 NQUTLPLAGLGLMHLKLGNALSOAFSKDSEPKLILEVPYAYOCCPYGMKASFEKA 491
DB 293 NQUTLPLAGLGLMHLKLGNALSOAFSKDSEPKLILEVPYAYOCCPYGMKASFEKA 352
QY 492 SGOWEADLHLDDESSKRPGLLARQENHYDODDELQLEMDSKPHPSVOCSPTRGP 551
DB 353 SGOWEADLHLDDESSKRPGLLARQENHYDODDELQLEMDSKPHPSVOCSPTRGP 412
QY 552 FKPEYLFESGIRLAWAIVLVSLVLCNGILVLFVAGGPAFLPYVGVGATAGANTLT 611
DB 413 FKPEYLFESGIRLAWAIVLVSLVLCNGILVLFVAGGPAFLPYVGVGATAGANTLT 472
QY 612 GISCGLLASVDALTFQGFSEYGARWETGLGCRANGFLAVIGSASVLLTLAAYQCSVS 671
DB 473 GISCGLLASVDALTFQGFSEYGARWETGLGCRANGFLAVIGSASVLLTLAAYQCSVS 532
QY 672 SCVRATGKSPISLGSVAGVIGCLATAGLAAALPLASVGEYSAPLCLPAPAPGGOAALG 731
DB 533 SCVRATGKSPISLGSVAGVIGCLATAGLAAALPLASVGEYSAPLCLPAPAPGGOAALG 592
QY 732 FTVAVLWMSFGLVAVAGVITKLYCDLPRGDEAVWDCAMVHVAMLLFADGLLYCPVAF 791
DB 593 FTVAVLWMSFGLVAVAGVITKLYCDLPRGDEAVWDCAMVHVAMLLFADGLLYCPVAF 652
QY 792 LSEAFMLGLFPYTPRAVSVLLVLDPLACINPLLYLFFNPHFRDLRLPRAGDSPL 851
DB 653 LSEAFMLGLFPYTPRAVSVLLVLDPLACINPLLYLFFNPHFRDLRLPRAGDSPL 712

QY 852 AYAAGELEKSSCDSTQALVAFSDVLTLEASSEAGRPGLTFYGPSPVTLISCOQPGAPR 911
DB 713 AYAAGELEKSSCDSTQALVAFSDVLTLEASSEAGRPGLTFYGPSPVTLISCOQPGAPR 772
QY 912 LEGSHCVPEGNHNGNPPSPMDGELLARABESTPAGGSLSGGGGFP 958
DB 773 LEGSHCVPEGNHNGNPPSPMDGELLARABESTPAGGSLSGGGGFP 819

RESULT 2

ID ABG09314 standard; Protein; 691 AA.

AC ABG09314;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9305.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

OS food supplement; medical imaging; diagnostic; genetic disorder.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PA 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS73501.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PS Claim 20; SEQ ID No 39673; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations in

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG30377 represent novel human

XX diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at http://wipo.int/publ/published_sequences.

SQ Sequence 691 AA:

Query Match 57.0%; Score 2886.5; DB 22; Length 691;
Best Local Similarity 89.1%; Pred. No. 1.4e-226;

Matches	580:	Conservative	10:	Mismatches	48:	Indels	13:	Gaps	9:
OY	316	MDIOEFPDLKGTSTLEITLIRAGIRLLPSGMCOOLPRVLYELSHNOIEELPSLRHCOK							375
Db	1	MDIOEFPDLKGTSTLEITLIRAGIRLLPSGMCOOLPRVLYELSHNOIEELPSLRHCOK							60
OY	376	LEETIGLOHNRIMEIGADTFQSLSLQALDSWNAIRSIHPEAFSTLSYKLDITDQLT							435
Db	61	LEETIGLOHNRIMEIGADTFQSLSLQALDSWNAIRSIHPEAFSTLSYKLDITDQLT							120
OY	436	TLPLAGLGLMHLKLGKLAISOAFSKSPFKRLIEPVYOCSPGMCSPFKASQOM							495
Db	121	TLPLAGLGLMHLKLGKLAISOAFSKSPFKRLIEPVYOCSPGMCSPFKASQOM							180
OY	496	EAEDHLHDEESSKRRPLGLARQAEHNYDQDDELQLEMEDSKRHPVSQSPRPGPKPC							555
Db	181	EAEDHLHDEESSKRRPLGLARQAEHNYDQDDELQLEMEDSKRHPVSQSPRPGPKPC							240
OY	556	EYLFESWGRILAVNAIVLISVLCNGLVLLTFAGGAPALPVKKFVYGAIAAGANTLTGISC							615
Db	241	EYLFESWGRILAVNAIVLISVLCNGLVLLTFAGGAPALPVKKFVYGAIAAGANTLTGISC							300
OY	616	GLLASVDALTFCGFSEYEG-ARWETELGCRATGFLAVLGSEASVLLTLTAAYQCSVSVSCV							674
Db	301	GLLASVDALTFCGFSEYEG-ARWETELGCRATGFLAVLGSEASVLLTLTAAYQCSVSVSCV							360
OY	675	RAYGKSPSLGSRVAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGAPALGFTV							734
Db	361	RAYGKSPSLGSRVAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGAPALGFTV							420
OY	735	ALYMNNSFCFLVAVGA---YIKLYEDLPGRDEAVWDCAMVHV--AM-LIFADGLLYCPV							789
Db	421	ALYMNNSFCFLVAVGA---YIKLYEDLPGRDEAVWDCAMVHV--AM-LIFADGLLYCPV							480
OY	790	AFLSFASMLGFLPVTPPEAVKSVLLVPLPACINPLLYLLFNPHRDDLRLLR-PRAGDS							848
Db	481	AFLSFASMLGFLPVTPPEAVKSVLLVPLPACINPLLYLLFNPHRDDLRLLR-PRAGDS							540
OY	849	GPLAYIAAGEL--EKSSCDS--TQALVAFSDYL--LLEASEA-GRPGLETGFPSTYLI							902
Db	541	GPLAYIAAGEL--EKSSCDS--TQALVAFSDYL--LLEASEA-GRPGLETGFPSTYLI							600
OY	903	SCQOPGAPRLGSHCVPEPGNHPGNPQPMDSGLLRARAGSTPAGGSLGG							953
Db	601	SCQOPGAPRLGSHCVPEPGNHPGNPQPMDSGLLRARAGSTPAGGSLGG							651

RESULT 3

AAM93889 standard; protein; 907 AA.

AAM93889.

25-JUN-1999 (first entry)

Human HG38 protein.

XX HG38: human; G-protein coupled glycoprotein hormone receptor; brain;
 XX endocrine system; skeletal muscle; spinal cord; placenta; development;
 XX receptor activity modulator.

XX Homo sapiens.

XX W09915660-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US19979.

XX 24-SEP-1997; 97US-0059863.

XX (MERI) MERCK & CO INC.

XX PA

PI	Bailey WJ, Liu Q, McDonald TP;								
XX	WPI, 1999-254711/21.								
DR	N-PSDB; AAM23960.								
XX	Human G-protein coupled glycoprotein hormone receptor HG38								
XX	Claim 1a; Fig 2; 74pp; English.								
CC	This invention describes a novel human G-protein coupled glycoprotein								
CC	hormone receptor, HG38. Glycoprotein hormone receptors are important in								
CC	the endocrine system and HG38 may be involved in development and function								
CC	of the skeletal muscle, spinal cord, placenta and to a lesser extent, the								
CC	brain. The transgenic animal may be useful for studying tissue and								
CC	temporal specific expression or activity of the HG38 receptor, as well as								
CC	for studying the ability of a variety of compounds to act as modulators								
CC	of HG38 receptor activity.								
XX									
SQ	Sequence 907 AA;								
Query Match	49.8%; Score 2525.5; DB 20; Length 907;								
Best Local Similarity	55.3%; Pred. No. 6e-197;								
Matches 497; Conservative 136; Mismatches 235; Indels 31; Gaps 9;									
OY	23 GGAPQPGPPTACAPCHQEDG-IMLSADCSSELGSAVBDLPLTAYLDLSMNNITEL								81
Db	22 GSSPRSGVLLRCGPTCHQCEPDGFMILLRYDCSDLGISELSPNLSVFTSYDLDSMNNISOL								81
OY	82 QPGFLFHLRFLFEELRLGNLSHTPGQAFSGIVSLKTIOMONNOGIPAEALMELPSIQ								141
Db	82 LPNPLPSLRFLDEELRLAGNALTYTPKGAFTGLYSLKVMLONNOLRHVPPEALONLSIQ								141
OY	142 SLRLDANLISVPRSEFGLSLRLHMLDNDALTEIVRALNNLPALQATLALNRISHI								201
Db	142 SLRLDANLISVPPSCSGSLRLHMLDNDALTEIVQAFRLSALQATLALNKIHI								201
OY	202 PDYAFQNTLSVLVHLNHNRIQHTSBSGLNLEFLDNYNKLOFFPAIRLIGLOE								261
Db	202 PDYAFGNLSLVVLIHNHNRIHSIGKCFDLSHLETLIDLYNNLDEFPPIRILSNLKE								261
OY	262 LGFHNHNKAIPEKAFMGNPLLOTTHFYDNPLOVVGSAFOYLRKLTLSINGAMDIOEF								321
Db	262 LGFHNHNRIPEKAFVGNPSLTIHFYDNPLOVVGSAFOYLRKLTLSINGAMDIOEF								321
OY	322 PDLKGTSTLEITLIRAGIRLLPSGMCOOLPRVLYELSHNOIEELPSLRHCOKLEITGL								381
Db	322 PDLTGTANDESLTLGAISSLPQTVYCNQPLNLOVLDSTYNLRLDPSFVCOKLQKIDL								381
OY	382 QHNRIMEIGADTFQSLSLQALDSWNAIRSIHPEAFSTLSYKLDITDQLTTLPLAG								441
Db	382 RHNEIYEIKVDTFOQLSLRSINLAWNKIATIHPEAFSTLSLRKLDLSSNLSLSPITG								441
OY	442 LGLMHLKLGKLAISOAFSKSPFKRLIEPVYOCSPGMCSPFKASQOM-----								495
Db	442 LHGLTHLKLGNHALOSLISSENFPELKIEMPRAYQCCAGVCEMAYKINOMNKGDS								501
OY	496 EAEDHLHDEESSKRRPLGLARQAEHNYDQDDELQLEMEDSKRHPVSQSPRPGPKPC								554
Db	502 SMDDH-----KKDAGMFOQODE-----RDLEDFLDFEEDLKALHSVQSPSPGPKPC								550
OY	555 CEYLFESWGRILAVNAIVLISVLCNGLVLLTFAGGAPALPVKKFVYGAIAAGANTLTGIS								614
Db	551 CEHLIDGWLIRIGVWTIALALTCNALVSTYER-SPLYSIPKILITGVIAVNMNLTGS								609
OY	615 CGLASVDALTFCGFSEYEGARWETELGCRATGFLAVLGSEASVLLTLTAAYQCSVSVSCV								674
Db	610 SAVIAGVDAFTFGSFARHGAMWENGCHVIGFLIFASSESVFLLTLAALERGVSYS								669
OY	675 RAYGKSPSLGSRVAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGAPALGFTV								734
Db	670 AKFETRAPPSLSKVIILLCALALATMAAVPLDGGSKYGASPLCLPL--PFGSPSTMGIMV								727
OY	735 ALYMNNSFCFLVAVGAYIKLYCDLPGRDEAVWDCAMVHVAMLIIFADGLLYCPVAFSLF								794

```

Db 728 ALILNLSLCEFLMTIATYKLYCNLDKGLDENIMDCSWKHIALLLPNCILNCPVAFLSF 787
OY 795 ASMLGLPPTPEAVKSVILVLPPLACLPNPLLYLLNPHFRDLRLRPPRA-----GDSG 849
Db 788 SSLINLTFISPEYKIFLLVWPPLACLPNPLLYLLNPHFRDLRLRPPRA-----GDSG 849
OY 850 PLAYAAAGELERKSSCDSTQALVAFSDVDLLLEASGRP-----PGLETYGFPSVTLLSC 904
Db 848 SLMSINSDVEKQSCDSTQALVAFSDVDLLLEASGRP-----PGLETYGFPSVTLLSC 906

RESULT 4
AA90682
ID AA90682 standard; Protein; 907 AA.
AC AA90682;
XX 21-AUG-2000 (first entry)
XX
XX Human G protein-coupled receptor HG38.
DE
XX
XX G protein-coupled receptor; GPCR: constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening;
XX agonist; antagonist.
XX
XX Homo sapiens.
OS
XX MO200022129-A1.
XX
XX 20-APR-2000.
XX
XX 12-OCT-1999; 99WO-US23938.
XX
XX 13-OCT-1998; 98US-0170496.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Behan DP, Chalmers DT, Liaw CW;
PI WPI; 2000-329165/28.
DR N-PSDB; AAA30770.
XX
XX Non-endogenous constitutively activated human G protein-coupled
PT receptor, useful for identifying agonists for use as pharmaceutical
PT agents -
XX
XX Example 1; Page 317-320; 341pp; English.
XX
XX The invention relates to constitutively active, non-endogenous versions
XX of endogenous human orphan G protein-coupled receptors (GPCRs, AA90643-
XX AA90677 and AA90683-190687), and to DNA encoding them (AAA30709-A30743
XX and AAA30775-A30779). The mutant proteins of the invention contain a
XX mutation in a portion of the protein comprising intracellular loop 3
XX (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
XX is substituted for an endogenous residue in IC3 at a position 16 amino
XX acids N-terminal of an endogenous proline in TM6 to form a sequence
XX X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
XX or Ala, and is preferably Lys. When the endogenous residue at this
XX position is Lys, this residue is replaced by His, Arg or preferably Ala.
XX The 15 amino acid stretch between the substituted amino acid and the Pro
XX may be endogenous, non-endogenous, or a mixture of endogenous and
XX non-endogenous residues. The constitutively active GPCRs are useful for
XX identifying antagonists, agonists and partially active GPCRs are useful for
XX pharmaceutical agents. The mutant proteins are also useful in research
XX settings for elucidating the roles of the receptors in normal and
XX diseased conditions. Antagonists for a particular GPCR are useful for
XX treating diseases and disorders associated with that receptor. Because
XX the novel mutant GPCRs are constitutively active, they can be used
XX directly for screening of compounds without the need for endogenous
XX ligands. The present sequence represents a human wild-type GPCR referred
XX to in an exemplification of the invention.

```

```

SQ Sequence 907 AA:
Query Match 49.8%; Score 2525.5; DB 21; Length 907;
Best Local Similarity 55.3%; Pred. No. 6e-197;
Matches 497; Conservative 136; Mismatches 235; Indels 31; Gaps 9;

OY 23 GGAPOGPPTACAPACQCOEDG-IMLSADCSLGLSAMPBGDLPPLAVILDSMNNITEL 81
Db 22 GSSPRGSGVLLRGCPTRHCHCEPDEGRMLLRVDCSDGLSELPSNLVSFTSYLSMNNISQL 81
OY 82 QPGLFHLRFLREELRLSGNHLSHIPGAFSGVSLKIMONNOGLGIPAEALMELPSIQ 141
Db 82 LRPRLPSLRFLEELRLAGNALITYPKGAFGLSLKVLMLONQHLHVPEALQNRSLQ 141
OY 142 SLRLDANLSLVERSEFEGSLSLRLMLDNLALTEIPVARNLNPALQAMTLALNRSHI 201
Db 142 SLRLDANHSIYVPPSCFSGLSLRLMLDNLALTEIPVAFSLSLQAMTLALNRSHI 201
OY 202 PDYAFQNLISLVVHLHNNRIQHLGTHSEFGLNLETLDLNNKLOEPFVAIRTLGRLOE 261
Db 202 PDYAFGNLSLVVHLHNNRIHSLGKCPDGLHSETTLDLNNNLDPEFTAIRTLNLKE 261
OY 262 LGFHNNNIKAIFPEKAFMGNDLQTIHFYDNPLOFVGSAFOYLPRLHTSLNGAMDOIEF 321
Db 262 LGFHSNNIRSIPEKAFVGNPSLITTHFYDNPLOFVGSAFOYLPRLHTSLNGASQITFE 321
OY 322 PDLKGTTSLEITLRAGIRLLPSGMCQOLPRLRVLELSHNOIEELPSLARCQKLEIGL 381
Db 322 PDLTSTANLESLLTGAQISSLPQTVCNQPLMLQVLDLSYNLLEDDPSFVSCQKLOKIDL 381
OY 382 QHNRIMEIGADPFSQSLSSLOALDLSMNAIRSHPEAFSTSLVYKLDLDNQLTLPLAG 441
Db 382 RHNEIYEIKVDYFQOLLISRLNLMANKIATIHDPAFSLPSLKLDDSSNLSSPITG 441
OY 442 LGLMLHLKLGMLALSOAFSKDSFPRKRLILEVRYAOCPPYGMCAFSKASGOW----- 495
Db 442 LGLHLKLTGNHALQSLISENPELKIEMPRYAYOCAPFCVENAKIKISQMKGONS 501
OY 496 EAEDLHDESSRRPLDLRLAQAENHYDQDLEQLDEM-EDSKHPVYOCSPPTPGPKP 554
Db 502 SMDLH-----KDAQMPQADE---RLEDELDFEEDLKLHVSQCSPPGPKP 550
OY 555 CEYLFEESWIGIRLAWAIVLLSVLCGLIVLIVFAGGAPLPPVPRVGAIGANTLTGIS 614
Db 551 CEHLIDGWLIRIGWTLAVLALTCALVTSYER-SPLIYSPKILLIVIAAVNMLTGS 609
OY 615 CGLLASVDALFFGQFSEYEGARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSCV 674
Db 610 SAVLAGVDAPFFGSPFARIGAMWENGVCYIGFLSIFSESSVFLLTALALRGFSYKXS 669
OY 675 RAYGKSPSLGSVRAGVIGCLALAGLAALPLASVGEAGSPICLPYAPPEQPALGFTV 734
Db 670 AKFETKAPFSLKATIIILCALALITMAAVPLLGSKSGKASPLCLPL--PREPPTMGIMV 727
OY 735 ALVMAISFCLVAVAGAVIKLYCDLPRGDPEAVWDCAMVRHAWILFADGLIYCPVAFLSF 794
Db 728 ALILNLSLCEFLMTIATYKLYCNLDKGLDENIMDCSWKHIALLLPNCILNCPVAFLSF 787
OY 795 ASMLGLPPTPEAVKSVILVLPPLACLPNPLLYLLNPHFRDLRLRPPRA-----GDSG 849
Db 788 SSLINLTFISPEYKIFLLVWPPLACLPNPLLYLLNPHFRDLRLRPPRA-----GDSG 849
OY 850 PLAYAAAGELERKSSCDSTQALVAFSDVDLLLEASGRP-----PGLETYGFPSVTLLSC 904
Db 848 SLMSINSDVEKQSCDSTQALVAFSDVDLLLEASGRP-----PGLETYGFPSVTLLSC 906

RESULT 5
AA90687
ID AA90687 standard; Protein; 907 AA.
XX
XX AA90687;
XX

```

[illegible]

QY	202	PDAFQNLISLVLLHNNRIQHLGTHSEEGAHNETLDLWNTKLOEPVPAIFRLIGLQE	261
Db	202	PDVAFGNLSSLVLLHNNRHSIGKCEDGLSETLDLNNLNDLPEAIRLSNKE	261
QY	262	LGFEHNNKRIPEKAMGNPLQTHFYFNPNLOFGRSAGFOYLPLKHLSTLNGAMDIOEF	321
Db	262	LGFEHNNKIRIPEKAVGNPSLITTHFYFNPIQFGRSAGFOHLPRLRLTLNGASQITEF	321
QY	322	PDLKGTSTLILTLTRAGIRLPLPSGCOOLPRLVLEISHNOIEELPSLRCKQLEIGL	381
Db	322	PDLTGANLSESLVLTGQNLSSLPQVYCNQDLPNLQVLDLSYNLEEDLPSFSYCKLOKIDL	381
QY	382	QHNRIMEIGADTSSQSLDQADLDSMNARLSIHPEAFSTLSHVLTLTDNQLTTLPIAG	441
Db	382	RHNEYIEKVDTEFOQLLSLTLNLANNKIIAIIHPNAFSLPLIKLIDSSNLSSFPITG	441
QY	442	LGGMHLKIKGNLALSGAFKDSFPKRILEVYAYOCPCYMCASFRRASGOW-----	495
Db	442	LHGTLHLKLGNAALGSLISENPELKVIEMPAAYQCCAFVCENAKYKISNMWKGDS	501
QY	496	EAEDILHDEDESSKRPGLLARQENHYDODLDELQLEM-EDSKPHPSYQCSPTTGPFRP	554
Db	502	SMDDLH-----KKDAGMFOAQDE----RDLEDFLDEDEEDLKALHSYQCSPPSGPKP	550
QY	555	CEYLFESGIGILAWAIVLLSYLCNGIYVLTGAGSAPLPLPVKIVVGAIGAGNTLTGIS	614
Db	551	CEHLIDGWLIRIGWTTAVLATLGNALVSTYVER-SPLYISPKLIGIYIAVNMVLTVS	609
QY	615	CGLASVALPLFFGOFSEYGAKMETGLGCRATGFLAVLAGEASVLLTTLAAVOCYSVSCV	674
Db	610	SAVLAGVDAFFGFSFARRHAGMEGVCCHVIGELISFASSESVFLLTLAALRGFSVYS	669
QY	675	RAYGSPSLGSRVAGVGLGCLAAALPLASVGEYSGPLCLPYAPRPGGAPALGFV	734
Db	670	AKFEKAPFSSILKVIITLLCALALATMAAVPLLGSKSYGASPLCLPL--PFGESPVMGIV	727
QY	735	ALVMMNSCFLVYVAGAYIKLYCDLPRGDPFAVMDCAVHRVAMLIFADGLLYCPVAFLSF	794
Db	728	ALILINSICFLMMTIAIYKLYCLNCDKGDLENIMDCSKKHIALLLTNCILNCPAFLSF	787
QY	795	ASMLGFLPVTPBEAVKSVLLVPLPLPACINPLLYLLENFHRDDLRLRPPA-----GDSG	849
Db	788	SSLINLTFISPEVIKFILLVVPPLPACINPLLYLLENFHRKEDLVSLRQTYVMTRSKHP	847
QY	850	PLATYAAAGELEKSSCDSTQALVAFSDVDLIENSAGNR---PGLETGFPSTVLISC	904
Db	848	SLMSINSDDVEKQSDSTQALVLTFTSSITYDLPSSVSPAPVIESCHLSVAFAVPC	906
RESULT 6			
ID	AAI53574	standard; Protein: 497 AA.	
XX	AAI53574;		
XX	AC		
XX	DT	15-FEB-2000 (first entry)	
XX	DE	Human gonadotropin receptor partial sequence #4.	
XX	KX	Human; gonadotropin receptor; screening; ligand; biomedical research;	
XX	KW	biochemical research; drug; hormone; reproductive tissue; infertility	
XX	KX	contraception.	
XX	OS	Homo sapiens.	
XX	PN	EP950711-A2.	
XX	PD	20-OCT-1999.	
XX	PF	02-FEB-1999; 99EP-0200303.	
XX	PR	06-FEB-1998; 98EP-0200357.	

KW anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;
 KW diabetes; endocrine system; skeletal muscle; spinal cord; placenta;
 KW transgenic animal.

XX Homo sapiens.

XX WO915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPI; 1999-254689/21.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Claim 2; Fig 2; 89pp; English.

CC This invention describes the isolation of a novel human G-protein coupled
 CC glycoprotein hormone receptor, AOMF05 which has anti-obesity and
 CC anti-diabetic activity. The AOMF05 gene, receptor and modulators are
 CC useful for treatment of pancreatic disease, obesity and diabetes.
 CC Glycoprotein hormone receptors are important in the endocrine system
 CC and AOMF05 may be involved in development and function of the skeletal
 CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
 CC transgenic animal may be useful for studying tissue and temporal
 CC specific expression or activity of the AOMF05 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of AOMF05 receptor activity.

CC Sequence 949 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 949;

Best Local Similarity 46.2%; Pred. No. 1.6e-156;

Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

QY 1 MSPPGRLALMTCALCASRRAGAPQPGPTACAPCHCEDGIMLSADSELSAV 60
 DB 1 MPPGRL-----LC--FLALGLGSGAGPGCAPPLCAAPSCGCD--RRVDSGKGLTAV 51
 QY 61 PGDDPLTAYLDLSNNNTLEQPGLEHLLRLELEIRLSGNHLSHIPGQAFSLYSKILM 120
 DB 52 PEGISAFQALDISNNITQLPEDAFKNPFLLEQLAGNDLSFTHPKALSGKELKVL 111
 QY 121 LONNOLGIPAEALWELSLQSLRDANLISVPERSEGLSLRLHMLDNLTEIPVR 180
 DB 112 LONNOLKTVPSAIRGSLALSRLDANHTSVPEDSFEGVLQHLMLDNLTEIPVR 171
 QY 181 ALNNLPALQAMTLALNRISHTPDYAFONTLSLVNHLNNRLOHGLTSEEGSLHLEETD 240
 DB 172 PLSNPTQALTLANKLSSIPDAFTLSLVNHLNNKTRISQCFDGLDLEETD 231
 QY 241 LNYNKLQEPVAIRTLRGLQELGFHNNNIIKAIPEKAFGNPLQTHFYDNPICQVGRSA 300
 DB 232 LNYNKLQEPVAIRTLRGLQELGFHNNNIIKAIPEKAFGNPLQTHFYDNPICQVGRSA 291
 QY 301 FOYLPKLHTLSLNGAMDQEPDLKGTSLLETLITRAGIRLLBSGMCQQLRRLVLELS 360
 DB 292 FHNLSLDSHSLVIRGASWQEPNLTGLVHLESLTLTGKISSIPNNLCOEQKMLTDLDS 351
 QY 361 HNOIEELPSLRCKLEIRIGLQHNRIWEIGADTFESQSLALDLDSNVAISIHPEARST 420
 DB 352 YNNIRDLPSFENGCHALEISIQRNQIYIKRGITQGLISLILDSKRLHEIHSRAEAT 411
 QY 421 LHSVVKLDLDNQTTLPLAGLGLMLKLGKLNALSGAFSKDSFPKRLILEVPAVQCC 480
 DB 412 LGRITNDVSNFELTSEFTGELNGLNOLKLVGNFKKALAAKDFVNLRSLSVPAVQCC 471

QY 481 PYGMCASFPAKSGWMEADLHLDDESSKRPGLIAQENHNDQDDELQLEMEDSKPH 540
 DB 472 AFWGCDSYANLNTENNLSLODSVAQKEGTADAAVNTSTLENEHSQT----- 518
 QY 541 PSVOCSEPTPGPFKCEYLFPSWGIKRLAVMAIVLLSVLCNGILVLTVPAGPAPLPYKVF 600
 DB 519 -IIHCTPSTGAFKCEYLLSSMIRLTVWFILVALLFNLLVITTTASCTSLPSSKLF 576
 QY 601 VQAIAGANTLTGISCGLASVDALTFQFSEYGARWETGLGCRATGFLAVGSEASVLL 660
 DB 577 IGLISVSNLMEGYITGLTLDVSWGRFAEFGIMWETGCGKVAAGFLAVFSSSAIFLL 636
 QY 661 TLAAGCSVSVCVRAGKSPSLGVSAGVGLGCLALAGLAALPLASVGEYGSPLCLPT 720
 DB 637 MLATVERSLSAKIDMKNGKSNHLKQFVALLAFLGATVAGCPPLFRGYSASPLCLPT 696
 QY 721 APPEGPPALGFTVALVMNNSFCFLVYAGAYIKLYCDLPBGDEAVWDCAMVRHVAMLIF 780
 DB 697 --PTGETPSLGLFTVTLVNLNSLAFILMAVITYTKLYCNLEKEDLSSENSQSMIKHVAMLIF 754
 QY 781 ADGLLYCPVAFLSPASMLGLEPVTPEAVKSVLLVLEPLACNPLDLYLFNPFRRD--- 837
 DB 755 TNCIFCPVAFESFADLITAISSPEIMKSVTLFPLACNPLVYVFPNPFKEDMKL 814
 QY 838 -LRLRPAGDSGLPVAAGAELK-----SSCDSTQALVAESVD--- 877
 DB 815 LKRRVTKSGSVSVSSISQGCLEDPFYDCGMYSHLQNLTVDCESFLTKRPVSCRH 874
 QY 878 LILEASEAGNPPGLFTYGPVSYTLISCOOP 907
 DB 875 LI-----KSHSPALAVASCQRP 892

RESULT 10

AAW93905
 ID AAW93905 standard; Protein; 949 AA.

AC AAW93905;

XX 29-JUN-1999 (first entry)

XX Human AOMF05 protein.

XX AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;

XX anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;

XX diabetes; endocrine system; skeletal muscle; spinal cord; placenta;

XX transgenic animal.

XX Homo sapiens.

XX WO915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPI; 1999-254689/21.

XX N-PSDB; AAX24012.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Claim 1; Fig 3A-F; 89pp; English.

XX This invention describes the isolation of a novel human G-protein coupled

XX glycoprotein hormone receptor, AOMF05 which has anti-obesity and

XX anti-diabetic activity. The AOMF05 gene, receptor and modulators are

CC useful for treatment of pancreatic disease, obesity and diabetes.
 CC Glycoprotein hormone receptors are important in the endocrine system
 CC and AOMF05 may be involved in development and function of the skeletal
 CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
 CC transgenic animal may be useful for studying tissue and temporal
 CC specific expression or activity of the AOMF05 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of AOMF05 receptor activity.

XX Sequence 949 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 949;
 Best Local Similarity 46.2%; Pred. No. 1.6e-156;
 Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

OY 1 MSPSPGLALMLCALCALSRAGAPOPGPPTACAPCHCEDEGIMLSADCELSGLSAY 60
    |||||
DB 1 MEGPLGL-----LC--FLALGLIGSAGPSGAPPLCAAPCSCDGD--RRVDCSGKGLTAV 51
    |||||
OY 61 PGDDLPITAYLDLSNNNTLTELQPGLEFHHRLFELEIRLSGNHLSHPQAFSGLYSLKILM 120
    |||||
DB 52 PEGLSAFQTQALDISNNNTITQLEDAFKNPFLEELQLAGNDLSFTHPKALSGLEKLYLT 111
    |||||
OY 121 LONNOLGIPALMELSLQSLRDANLISLYPERSFEGLSLRHMLDNDALTEIPVR 180
    |||||
DB 112 LONNOLKTVPSFAIRGLSALQSLRDANHTSVPEDSFEGLYQRLHMLDNDNSLFEVPH 171
    |||||
OY 181 ALNNLPALQAMTLALNRISHIIPDYAFQNLTLVLVHLHNNRIQHLGTHSFEGLHNLFTLD 240
    |||||
DB 172 PLSNPLTQALTLALNKISSIPDFAFTNLSLVHLHNNKIRLSQCFQDGLDMLLETILD 231
    |||||
OY 241 LNYNKLQEPVPAIRTLQRLQELGFHNNNTKAIPEKAFKGNPLTQTHYDNPQIVGSRSA 300
    |||||
DB 232 LNYNNGEFPQAIKALPSLKELEGFHSNSISVIPDAFQDNPLRLTHIYDNPVGSNA 291
    |||||
OY 301 FQYLPRLHTLSLNGAMDIOEPDLKGTSTLETLTLRAGIRLPSGMCQQLRLVLELS 360
    |||||
DB 292 FHNLSLHSLVLRGASMVQAFRLTGTVHESTLTLTGTSISSIPNNLCQEQKMLRTILDS 351
    |||||
OY 361 HNOIEELPSLHRCQLEELIGLOHNRITWETGADTFQSLSQALDLSWNAIRSIHEPAFT 420
    |||||
DB 352 YNNIRDLPSNGCHALEEISLQNRQIYQKEGFGGLISRLIDYSRNLIHEHSAFAYT 411
    |||||
OY 421 LHSVLVLDLTNQLTLPLAGLGGMLHKLKGLALSOAFKSDKFLILEPVAAYQCC 480
    |||||
DB 412 LGRITNLVDSENFELTFPEEGMLGQLKLVGNFKLEKALAKDFVNLISLSPVAYQCC 471
    |||||
OY 481 PYGCASFPAAGOMEADLHLDDESSKRPLGLLAROENHYDODLDELQLEMEDSKPH 540
    |||||
DB 472 AFNGCDSYALNNTENNLSIDHSAQEGTADANVTSTLENEHSQI----- 518
    |||||
OY 541 PSVOCSETPGPPEKPCETLFEFSWGRILAVNAVILSVLCNGLVLLVYFAGGPAFLPVKTY 600
    |||||
DB 519 -IIHCTPSTGAFKPCETELGSSMMIRLTWFIYVLFENLITLTTFASCT--LPESKLEP 576
    |||||
OY 601 VGAIAGANTLTGSCGLLAVDALTFEGFSEYARWETGICGRATFLAVLGSEASVLL 660
    |||||
DB 577 IGLISVNLFWGITYTGLTFLDAVSNGRFAPFGIMWETSGCVAAGLAFSSESAIFLL 636
    |||||
OY 661 TLAAVQCSVSVCYRAYGKPSLGSVAVGLGIALAALPLASVGEYAGSPICLFLP 720
    |||||
DB 637 MLAVVERSLSKOIMKNGSKNHLKQFRVAAALFLGATVAGCPFLFERGETYSASPLCLPF 696
    |||||
OY 721 APPGQRAALGFTYALVYVMSFCFLVYVAGVYIKLYCDLPBGDFEAVYDCAWVRVAVLIF 780
    |||||
DB 697 --PGTEPSTLGFYTLVLLNSLAEFLMAVITYTKLYCNLEKEDLSSENSOSSNIKRVAMLIF 754
    |||||
OY 781 ADGLIYCVPAFLPSASMLGEPVTPPEAVKSVTLVVLPLPACLNPLLYLTFNPHFRD-- 837
    |||||
DB 755 TNCIFPFCVAFAPPLITATISIPELIMKSTVTLFFPLPACLNPLVLYVFEFNPKEKEDMKL 814
    |||||
OY 838 -LRLRPRAAGSGPLAVAAAGELEK-----SSCDSTQALVAFADVD-- 877
    |||||

```

DB 815 LKRRVTKSGSVSVSISQGGCLEDEYYDCMGYSHLOGNLTVCDCBSFLLTKPEVSCNH 874

OY 878 LTIASEAGRPPELETGYRPPSYTLISCCQP 907

DB 875 LI-----KSHSCPALAVASCORP 892

RESULT 11

AAW93965 ID AAW93965 standard; Protein; 951 AA.

XX AAW93965;

XX 29-JUN-1999 (first entry)

XX Human AOMF05 protein.

XX AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;

XX anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;

XX diabetes; endocrine system; skeletal muscle; spinal cord; placenta;

XX transgenic animal.

XX Homo sapiens.

XX WO9915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPI: 1999-254689/21.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Disclosure; Fig 5; 89pp; English.

This invention describes the isolation of a novel human G-protein coupled glycoprotein hormone receptor, AOMF05 which has anti-obesity and anti-diabetic activity. The AOMF05 gene, receptor and modulators are useful for treatment of pancreatic disease, obesity and diabetes. Glycoprotein hormone receptors are important in the endocrine system and AOMF05 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the AOMF05 receptor, as well as for studying the ability of a variety of compounds to act as modulators of AOMF05 receptor activity.

XX Sequence 951 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 951;
 Best Local Similarity 46.2%; Pred. No. 1.6e-156;
 Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

OY 1 MSPSPGLALMLCALCALSRAGAPOPGPPTACAPCHCEDEGIMLSADCELSGLSAY 60
    |||||
DB 1 MEGPLGL-----LC--FLALGLIGSAGPSGAPPLCAAPCSCDGD--RRVDCSGKGLTAV 51
    |||||
OY 61 PGDDLPITAYLDLSNNNTLTELQPGLEFHHRLFELEIRLSGNHLSHPQAFSGLYSLKILM 120
    |||||
DB 52 PEGLSAFQTQALDISNNNTITQLEDAFKNPFLEELQLAGNDLSFTHPKALSGLEKLYLT 111
    |||||
OY 121 LONNOLGIPALMELSLQSLRDANLISLYPERSFEGLSLRHMLDNDALTEIPVR 180
    |||||
DB 112 LONNOLKTVPSFAIRGLSALQSLRDANHTSVPEDSFEGLYQRLHMLDNDNSLFEVPH 171
    |||||
OY 181 ALNNLPALQAMTLALNRISHIIPDYAFQNLTLVLVHLHNNRIQHLGTHSFEGLHNLFTLD 240
    |||||

```

```

Db 172 PLSNLPLOALTLALNKISSIPDFAFTNLSIVLHANNKISLSQHCDFGLDNETLD 231
QY 241 LNNKXLOEPPVAIRTLGRLOELGFHNNNNKAIPEKAFMGNPLQTHFYDNPLOFVGSA 300
Db 232 LNNNGEPPQAIKALPSIKELGFHNSISIVIPDGAFDGNPLRTIHLNDNPISFVNSA 291
QY 301 FOYLPKLTLSLNGAMDIQDFPDLKGTSTSEILTLTRAGIRLLPSGCOQLPRLVLELS 360
Db 292 FHNLSDLHSIVIRGASWVOQFPMLTGVHLESTLTGTKISSIPNNLCQEQKMLRTLDLS 351
QY 361 HNOIEPLSLHRCQKLEELGLOHNRIMEIGADTFQSLSLOALDLSMNAIRSHPEAFST 420
Db 352 YNNIRDLPSFNGCHALEEISLQRNQIYQKEGTFQGLISIRLIDVSRNLIHEHSAFAT 411
QY 421 LHSIVKLDLTDNOLTLPLAGLGMHLKLGMLALSOAFSKDSFPLKILEVAYQCC 480
Db 412 LGPTNLVDVSNELTSFPEGLNGLNQKLGVNFKLEKLAARDFVNLRLSLSPVAYQCC 471
QY 481 PYGCASFPPKASGQWAEEDLHLDDEESSKRPGLGLARQAEHNDYDQDDELQLEMEDSKPH 540
Db 472 AFWGCDSYANLNTENNISLODHSVAQEKGTADAANVTSTLENEHSQI----- 518
QY 541 PSVOCSPPTGPPKPCETYLEESGIRLAWAIVLVSVCGLVLLTVFAGGPAPLPVKFV 600
Db 519 -IIHCTPSTGAFKPCETYLEESGIRLAWAIVLVSVCGLVLLTVFAGGPAPLPVKFV 600
QY 601 VGLAIGANTLTGISCGLASVADALTFQGFSEYGAEMETGLGRATGFLAVLSEASVLL 660
Db 577 IGLISVSNLMFMGIYGLITFLDAVSMGRFAEFGIMMETSGCKVAGFLAVSESAFILL 636
QY 661 TLAAGVGSVSCVAVAKSPSLGVRAGVIGCLALAGLAAALPLASVGEYASPLCLPY 720
Db 637 MLATYERLSAKDKIMKNSHNLKQFVAALLAFLGATVAGCGPFLFHGEYSASPLCLPF 696
QY 721 APPEGPALATFTVALVMANSEFELVAGAYIKIKCDLPBGFEAVMWCAMRHANLIF 780
Db 697 --PGEFSPISGFTVTLVNLNLAFLMAVITKLCNLEKEDLSNOSSMKHAAMLIF 754
QY 781 ADGLXCVAVLFRASMLGLFPVPEAVKSVLVLPACLNPLLYLLFNPFRD--- 837
Db 755 TNCFFCQVAFSFAPLITTAISPEIKMSVTLTFPLPACLNPLVLYVFNPKREDMKL 814
QY 838 -LRLRLPRAGSGPLAVAAAGLEK-----SSCDSTQALAFSYVD--- 877
Db 815 LKRVRTKSSGSVSISISOGGCLQDPRYDQGMSHLDGNLTLVCDCCESFLLTRPVSCKH 874
QY 878 LILEASAGRPGLTETGFPSTVLISCOOP 907
Db 875 LI-----KSHSCPALAVASCORP 892

```

RESULT 12

AAW93906
ID AAW93906 standard; Protein; 951 AA.

AAW93906;

29-JUN-1999 (first entry)

Human AOMF05 protein.

AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain; anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity; diabetes; endocrine system; skeletal muscle; spinal cord; placenta; transgenic animal.

OS Homo sapiens.

XX WO9915545-A1.

PD 01-APR-1999.

XX

```

PF 24-SEP-1998; 98WO-0520101.
XX
PR 24-SEP-1997; 97US-0059868.
XX
PA (MERT) MERCK & CO INC.
XX
PI Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;
XX WPI; 1999-254689/21.
XX N-PSDB; AAX24014.
XX
PS Human G-protein coupled glycoprotein hormone receptor AOMF05
XX disclosure; Fig 6A-F; 89pp; English.
XX
CC This invention describes the isolation of a novel human G-protein coupled
CC glycoprotein hormone receptor, AOMF05 which has anti-obesity and
CC anti-diabetic activity. The AOMF05 gene, receptor and modulators are
CC useful for treatment of pancreatic disease, obesity and diabetes.
CC Glycoprotein hormone receptors are important in the endocrine system
CC and AOMF05 may be involved in development and function of the skeletal
CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
CC transgenic animal may be useful for studying tissue and temporal
CC specific expression or activity of the AOMF05 receptor, as well as
CC for studying the ability of a variety of compounds to act as modulators
CC of AOMF05 receptor activity.
CC
XX
SQ Sequence 951 AA:

```

Query Match 40.1%; Score 2030.5; DB 20; Length 951;
Best Local Similarity 46.2%; Pred. No. 1.6e-156;
Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

QY 1 MSPPGRLAMTICAAALCASRRAGAPQPPGPTACPARCQCEDEIMLSADCEISGLSAV 60
Db 1 MEGPLGL-----LC--FLALGLGSAGPSGAAPPLCAAPCCSGD--RRVDSGGGLAV 51
QY 61 PGDDPLTAVYLDLSMNNLTTELQPLFHLRLFEELRLSGNHLSHPGAFSGLSKLTM 120
Db 52 PEGLSAFQALDISNNITQLPEDAFKFPLEELQTLGNDLSFTHRPALGLKELKYL 111
QY 121 LONNOLGIPALMELPLOSILRDLANLISVPERSEGLSLRLHMLWDNALTEIVPR 180
Db 112 LONNOLKTVPESEAIRGSLQSLRDLANHITSVPDESGLYQRLHMLDONSLEIYVH 171
QY 181 ALNNLPALQAMTLALNRTISHIDYAFQNLTSIVYLHLNNKTIQHLGHSFEGLANLETLD 240
Db 172 PLSNLPLOALTLALNKISSIPDFAFTNLSIVLHANNKISLSQHCDFGLDNETLD 231
QY 241 LNNKXLOEPPVAIRTLGRLOELGFHNNNNKAIPEKAFMGNPLQTHFYDNPLOFVGSA 300
Db 232 LNNNGEPPQAIKALPSIKELGFHNSISIVIPDGAFDGNPLRTIHLNDNPISFVNSA 291
QY 301 FOYLPKLTLSLNGAMDIQDFPDLKGTSTSEILTLTRAGIRLLPSGCOQLPRLVLELS 360
Db 292 FHNLSDLHSIVIRGASWVOQFPMLTGVHLESTLTGTKISSIPNNLCQEQKMLRTLDLS 351
QY 361 HNOIEPLSLHRCQKLEELGLOHNRIMEIGADTFQSLSLOALDLSMNAIRSHPEAFST 420
Db 352 YNNIRDLPSFNGCHALEEISLQRNQIYQKEGTFQGLISIRLIDVSRNLIHEHSAFAT 411
QY 421 LHSIVKLDLTDNOLTLPLAGLGMHLKLGMLALSOAFSKDSFPLKILEVAYQCC 480
Db 412 LGPTNLVDVSNELTSFPEGLNGLNQKLGVNFKLEKLAARDFVNLRLSLSPVAYQCC 471
QY 481 PYGCASFPPKASGQWAEEDLHLDDEESSKRPGLGLARQAEHNDYDQDDELQLEMEDSKPH 540
Db 472 AFWGCDSYANLNTENNISLODHSVAQEKGTADAANVTSTLENEHSQI----- 518
QY 541 PSVOCSPPTGPPKPCETYLEESGIRLAWAIVLVSVCGLVLLTVFAGGPAPLPVKFV 600
Db 519 -IIHCTPSTGAFKPCETYLEESGIRLAWAIVLVSVCGLVLLTVFAGGPAPLPVKFV 600

```

[illegible]

XX 24-SEP-1997; 97US-0059863.
 XX (MERI) MERCK & CO INC.
 XX PI Bailey WJ, Liu Q, McDonald TP;
 XX WPI: 1999-254711/21.
 DR N-PSDB: AAX23981.

XX Human G-protein coupled glycoprotein hormone receptor HG38

XX PS Disclosure: Fig 3A-E; 74pp: English.

CC This invention describes a novel human G-protein coupled glycoprotein
 CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
 CC the endocrine system and HG38 may be involved in development and function
 CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
 CC brain. The transgenic animal may be useful for studying tissue and
 CC temporal specific expression or activity of the HG38 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of HG38 receptor activity.

XX SQ Sequence 692 AA;

Query Match 39.4%; Score 1996; DB 20; Length 692;

Best Local Similarity 57.9%; Pred. No. 6.7e-154;

Matches 393; Conservative 95; Mismatches 171; Indels 20; Gaps 6;

OY 23 GGAPQPGPPTACAPACHOEDG-IMLSADCSRLGSAVPGDLDPLTAYIDLSMNNLTSL 81
 DB 22 GSSPRSGVLLRCGPTHCHCEPDGRMLRYDCSDGLSELPSNLSVTSTYIDLSMNNISQL 81
 OY 82 QPGLFHLHLFLELRLSGNHLSHIPQAFSGYSLKILMLQNNQGLGIPAEALMELPSIQ 141
 DB 82 LNPPLPSLFLFELRLAGNALTYIPKGAFTGYSLKVLMLQNNQRLHVPTEALQNLRSIQ 141
 OY 142 SLRLDANLISLPERFEGLSLRHMLDNDNATEIPVALNPLPALQAMTLALNRISHI 201
 DB 142 SLRLDANLISLPERFEGLSLRHMLDNDNATEIPVALNPLPALQAMTLALNRISHI 201
 OY 202 PDYAFONTLSLVYHLHNNRIQHLGTHSEFGLNLETLDNYNKLOEFPVARTLGRLOE 261
 DB 202 PDYAFONTLSLVYHLHNNRIQHLGTHSEFGLNLETLDNYNKLOEFPVARTLGRLOE 261
 OY 262 LGFHNNNIKAIPKAFWGNPLQTHFYDNPIDFVGSFAFYLPKLTLSLNGAMDIOEF 321
 DB 262 LGFHNNNIKAIPKAFWGNPLQTHFYDNPIDFVGSFAFYLPKLTLSLNGAMDIOEF 321
 OY 322 PDLAGTTSLEILLTRAGITLLPSGMOQLPRLRVLELSHNOIEELPSLRCKLEIGL 381
 DB 322 PDLAGTTSLEILLTRAGITLLPSGMOQLPRLRVLELSHNOIEELPSLRCKLEIGL 381
 OY 382 QHNRIMEIGADTFESQSLQALDLSWNAISIHPEAFSTLSLVKLDLTDNQLTTLPLAG 441
 DB 382 QHNRIMEIGADTFESQSLQALDLSWNAISIHPEAFSTLSLVKLDLTDNQLTTLPLAG 441
 OY 442 LGLMLHLKLGKALALSOAFKSDSPKRLILEVPRAYOCCPYGMCAFFKASGOW----- 495
 DB 442 LGLMLHLKLGKALALSOAFKSDSPKRLILEVPRAYOCCPYGMCAFFKASGOW----- 495
 OY 496 EAEDELHDDDESSSRPLGLLARQENHYDDLDLEOLEM-EDSKRHPSPVOCSPTPGPFPK 554
 DB 496 EAEDELHDDDESSSRPLGLLARQENHYDDLDLEOLEM-EDSKRHPSPVOCSPTPGPFPK 554
 OY 502 SMDDLH-----KKDAGMFOAQDE----RLDEFLIDFEEDLKALHSVQCSPPGPFKP 550
 DB 502 SMDDLH-----KKDAGMFOAQDE----RLDEFLIDFEEDLKALHSVQCSPPGPFKP 550
 OY 555 CEYLFEWSGIRLAWAIVLVLCNGVLVLTFRAGGAPALDPVKFVVGAIGANTLIGIS 614
 DB 555 CEYLFEWSGIRLAWAIVLVLCNGVLVLTFRAGGAPALDPVKFVVGAIGANTLIGIS 614
 OY 551 CERHLDGMLIRIGVWTIAVLATCNALVTSTVFR-SPLYISPIKILIGVIAVNMLTGVS 609
 DB 551 CERHLDGMLIRIGVWTIAVLATCNALVTSTVFR-SPLYISPIKILIGVIAVNMLTGVS 609
 OY 615 CGLASVDALTFEGQSEVGARWETIGCRATGFLAVGSEASVLLTLTAOVCSVSCV 674
 DB 615 CGLASVDALTFEGQSEVGARWETIGCRATGFLAVGSEASVLLTLTAOVCSVSCV 674
 OY 610 SAVIAGVDAFTFGSFARHGAWMENGVCCHVTIGLSIFRASESSVFLITLALERGFVAKYS 669
 DB 610 SAVIAGVDAFTFGSFARHGAWMENGVCCHVTIGLSIFRASESSVFLITLALERGFVAKYS 669

OY 675 RAYGKSPSLGSVRAGVLGC 693
 DB 670 AKFTKAPSSILKVIILLC 698

Search completed: October 23, 2002, 10:43:27
 Job time : 43 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52 ; Search time 26 Seconds
(without alignments)
3573.784 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MSPSPGIRALMLCALICASR.....GGLSGGGGFGPSPGIAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*

1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2525.5	49.8	907	JE0176	orphan G protein-c
2	2443.5	48.2	907	JC0193	foliitropin-recept
3	534	10.5	694	JC4301	foliitropin-recept
4	526.5	10.4	695	I45896	foliitropin-recept
5	525	10.4	694	JC2237	foliitropin-recept
6	524.5	10.3	695	JC1493	foliitropin-recept
7	519.5	10.2	692	A34548	foliitropin-recept
8	518	10.2	695	ORHUT	foliitropin-recept
9	515	10.2	695	ORH898	foliitropin-recept
10	490.5	9.7	764	A40077	thyrotropin-recept
11	486	9.6	764	I48882	thyrotropin-recept
12	484.5	9.6	1091	A58532	glial cell membran
13	476	9.4	814	JC7389	thyroid stimulat
14	474	9.4	700	I77463	luteinizing hormon
15	474	9.4	700	A49744	luteinizing hormon
16	473.5	9.3	764	JC3643	thyroid stimulat
17	472.5	9.3	793	JC7390	thyroid stimulat
18	472	9.3	700	A42395	thyrotropin-recept
19	469.5	9.3	764	ORHURH	thyrotropin-recept
20	467	9.2	696	A41344	luteinizing hormon
21	463	9.1	696	JC7361	foliitropin-recept
22	462.5	9.1	764	A35956	thyrotropin-recept
23	455.5	9.0	693	JC1282	insulin-like growt
24	439.5	8.7	699	ORHUT	insulin-like growt
25	439	8.7	695	JC5239	insulin-like growt
26	438.5	8.7	603	JC6128	insulin-like growt
27	437.5	8.6	605	A41915	insulin-like growt
28	432.5	8.5	889	T20123	hypothetical prote
29	411	8.1	925	JC2033	G protein-coupled

30	397	7.8	1469	2	B36655	silt protein 2 pre
31	397	7.8	1480	2	A36655	silt protein 1 pre
32	394.5	7.8	1531	2	T42218	silt-1 protein hom
33	384.5	7.6	1066	2	T15864	hypothetical prote
34	384	7.6	1385	2	T13887	tlr protein - frui
35	383.5	7.5	1523	2	T13953	MGEF5 protein - ra
36	382	7.5	1389	2	T13852	gene wheeler prote
37	375	7.4	1389	2	A60164	platelet membrane
38	375	7.4	1119	2	AD1822	leucine-rich-repea
39	364.5	7.2	1134	1	A29944	chaoptin precursor
40	363	7.2	536	2	A34901	lysine carboxypept
41	363	7.2	662	2	S42799	gap precursor
42	360.5	7.1	738	2	T19938	hypothetical prote
43	354.5	7.0	961	2	T23395	hypothetical prote
44	348.5	6.9	1115	2	S40241	G protein-coupled
45	340.5	6.7	1039	2	T22117	hypothetical prote

ALIGNMENTS

RESULT 1

JE0176
orphan G protein-coupled receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C:Accession: JE0176
R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
A:Title: Identification and cloning of an orphan G protein-coupled receptor of the gl
A:Reference number: JE0176; M0ID:98508104
A:Accession: JE0176
A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: GB:AF062005; NID:q3366801; PIDN:AAC28019.1; PID:q3366802
C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:639-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match	49.8%	Score 2525.5	DB 2	Length 907
Best Local Similarity	55.3%	Pred. No. 1.2e-168		
Matches 497	Conservative 136	Mismatches 235	Indels 31	Gaps 9
QY	23	GGAGPQPGPTACPAFCOEDG-IMLSADCESEGLSAVFGDDLPATVAYIDLSNNITEL	81	
DB	22	GSSRSVGLRGCTHCHCEPDGMLRVCOSDGLSEPSNLVSFTSYDLSNNITSL	81	
QY	82	QPGLEHRLFELEIRLSGNHLSHPGAFSGLSLKLIMQNNQDGLGPRMALWELSLQ	141	
DB	82	LPNPLPSLRFLELEIRLSGNHLSHPGAFSGLSLKLIMQNNQDGLGPRMALWELSLQ	141	
QY	142	SLRLDANLISVPRSEFGLSLRLHMLDNDNALTEIPRALNIPALQAMTLALNRISHI	201	
DB	142	SLRLDANLISVPRSEFGLSLRLHMLDNDNALTEIPRALNIPALQAMTLALNRISHI	201	
QY	202	PDVAFQNLTSVLVHLHNNRIQHLGTHSEGLHNLFTLDLNYKLOEPVPAIRLGRLOE	261	
DB	202	PDVAFQNLTSVLVHLHNNRIQHLGTHSEGLHNLFTLDLNYKLOEPVPAIRLGRLOE	261	
QY	262	LGFINNNIKALPEKAFKQNPLOTIHFYDNPLOFGREAFQYLPKRLTSLSGANDIOEF	321	
DB	262	LGFINNNIKALPEKAFKQNPLOTIHFYDNPLOFGREAFQYLPKRLTSLSGANDIOEF	321	
QY	322	PDLKGTSLLEITLITLRAGIRLLPSGMCOOLPRLRVYLSELHNOIEELPSLRQCUIEIGL	381	

```

      322 PDLGTANLESILTLTGAQSSLPQIVCNQDLPNQLVDLSLNLEDDLPSPSVCKKQKIDL 381
QY 382 QHNRIWEIGADPFSSQSLQALDLSMNAIRSHPEAFSTLSLVKLDLTDNQLTTLPLAG 441
Db 382 RINEIYEIVDFQQLSLRSLNLANWKTAIHPNAFSTLPLIKDLSSNLSLSPITG 441
QY 442 LGLMLHLKLGKLNALSOAFSKDSEPKRLILEVPAVQCCPYGMCASFRRASGQW----- 495
Db 442 LHLGLTHLKLGTGNHALQSLSSSEFPFLKTIEMPAIYQCCAFVCENAYKISNOMKNGDS 501
QY 496 EAEEDLHLDDESSKRPGLGLARQAEHNYDQDLDELQLEM-EDSKPHPSVQCSPPGPFKRP 554
Db 502 SMDDH-----KQDAGMFOAQDE---RDEDFLLDDEEDLKALHSVQCSPPGPFKRP 550
QY 555 CEYLESMGIRLAVMAIVLVLNGLVLTVPFAGAPALPVPKFEVVGAIAGANTLTGIS 614
Db 551 CEHLIDGMLIRIGVWTIAVLALCNALVTIVFR-SPLYSIPKILGIYIAAVNMMLTGV 609
QY 615 GGLLASVDALTFGQSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSCV 674
Db 610 SAVLAGVDAFTFGSFARHAGAMWENGCGVIGFLSTFASSESVFLTLTALBERGSVKYS 669
QY 675 RAYGSPSLGSRVAVLGLALAGLAALPLASVEYSGASPLCLPYAPPEGOPALGFTV 734
Db 670 AKFETKAPFSSLKVTILLCALALTMMAVPLLGSGSKYGASPLCLPL--PFGPSTMGWYV 727
QY 735 ALVMNSCFCLVAVAGATYIKLYCDLPRGDFEAVDCAMVRHVMILFADGLLYCPAFISF 794
Db 728 ALIILNSICFLMTATYIKLYCNLDKGLDENTMDCSMVKHIALTLFTNCILCPAFISF 787
QY 795 ASMLGLFVTPBAVKSLLVVLPLPACLNPLLYLLENPHFRDRLRLPRA-----GDSG 849
Db 788 SSLINLTPISEVIFILLVVPPLPACLNPLLYLLENPHFRDRLRLPRA-----GDSG 849
QY 850 PLAAVAGELKSSDSTQALVAFSDVDLLEASAGRP-----PGLETYGPPSVTLISC 904
Db 848 SLMSINSDVERKQSDSTQALVFTSSITYDLPPSVSPAYPTESCHLSVAFVPC 906

RESULT 2
G protein-coupled receptor FEX - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Jul-1999 #sequence.revision 23-Jul-1999 #text.change 11-May-2000
C/Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A>Title: Identification of a novel seven-transmembrane receptor with homology to glycoprotein A; Accession: JG0193; MUID:99121227
A:Reference number: JG0193; MUID:99121227
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>

Query Match 48.2%; Score 2443.5; DB 2; Length 907;
Best Local Similarity 54.3%; Pred. No. 6.9e-163;
Matches 485; Conservative 146; Mismatches 244; Indels 19; Gaps 9;

QY 22 AGGAPGPGPTACPAFCHQEDG-IMLSADCSFELGSAVPGDDPLVATYLDLSMNNITE 80
Db 21 AGSPGPDALPRGSPCHCELDGRMLRLKRDGSLGSELPSNLSVTSTYDLSMNNISQ 80
QY 81 LQGLPLHLFLDELRLSGNHLSHIPEQAFSGLYSKITLMQNNQGLGIPAAELWELP 140
Db 81 LPASLHRLCFLEELRLAGNALTHIPKAFGTGHSKLVLMQNNQKRVDEALQNRSL 140
QY 141 OSRLDANLISLVPERSEFELSLRHLMDNMLTEIPVALANLPRALQAMTALANISH 200
Db 141 OSRLDANLISLVPPSCFSLSRHLMDNMLTDVPAVAFSLSLQAMTALANISH 200
QY 201 IDPYAFQNTLSLVVHLHNNRIQHLGTHSPEGLHNLTDLDLNNKLOEFPVATRTLRQ 260
Db 201 IDPYAFQNTLSLVVHLHNNRIQHLGTHSPEGLHNLTDLDLNNKLOEFPVATRTLRQ 260

```

```

Db 201 IDPYAFQNTLSLVVHLHNNRIHSLKCPDGLHSETTLDLNNNLEDEPTAIKTLNLK 260
QY 261 ELGFHNNNIKAPEKAFMGNPLQTHFYDNPIQFGRSAFOYLPKRLHTLSLGNADIOE 320
Db 261 ELGFHNNNIKSPERAFVGNPSLITTHFYDNPLOFVGSFAFOHLPKRLHTLSLGNADIOE 320
QY 321 FPDKSTSLLEILLTRAGIRLLPSGMOQLPRLRYLESHNOIEELPSLHRCOKLEIG 380
Db 321 FPHLTGTATLESLTLTGAKISSLPQAVCDQLPMLQVLDLSYNLEDDLPSSGCKQKIDL 380
QY 381 LOHNRIMEIGADPFSSQSLQALDLSMNAIRSHPEAFSTLSLVKLDLTDNQLTTLPLA 440
Db 381 RINEIYEIVDFQQLSLRSLNLANWKTAIHPNAFSTLPLIKDLSSNLSLSPITG 440
QY 441 GLGLMLHLKLGKLNALSOAFSKDSEPKRLILEVPAVQCCPYGMCASFRRASGQWAEED 500
Db 441 GLHGLTHLKLGTGNHALQSLSSSEFPFLKTIEMPAIYQCCAFVCENAYKISNOMKNGD 500
QY 501 HLDDESSKRPGLGLARQAEHNYDQDLDELQLEM-EDSKPHPSVQCSPPGPFKPEYLF 559
Db 501 NSVD-DLHKKDAGLRQVODE---RDEDFLLDDEEDLNALHSVQCSPPGPFKPEHLE 555
QY 560 ESMGIRLAVMAIVLVLNGLVLTVPFAGAPALPVPKFEVVGAIAGANTLTGISGLLA 619
Db 556 GSWLIRIGVWTIAVLISCNALVATVFR-TPYISSLIKLLGVTAVDIIMGVSSAVLA 614
QY 620 SYDALTFGQSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSCVAYK 679
Db 615 AYDAFTFGFRAHOGAMWEDIGCQIVGFLSTFASSESVFLTLTALBERGSVKSEY 674
QY 680 SPISGSRVAVLGLALAGLAALPLASVEYSGASPLCLPYAPPEGOPALGFTV 739
Db 675 KAPLFSRLRIYLVCLVALLTLTIPLLGSKYNASPLCLPL--PGEESTGTGMVALVLL 732
QY 740 NSFCELVAVAGATYIKLYCDLPRGDFEAVDCAMVRHVMILFADGLLYCPAFISF 799
Db 733 NSICPLMTIATYIKLYCNLEKELELMDCSMVKHIALTLFTNCILCPAFISF 792
QY 800 LEPTPEAVKSVLVLPLPACLNPLLYLLENPHFRDRLRLPRA-----GDSGPLAVA 854
Db 793 LTFSPDVIKFLLVIVPLPCLNPLLYLLENPHFRDRLRLPRA-----GDSGPLAVA 852
QY 855 AAGELEKSSDSTQALVAFSDVDLLEASAGRP-----PGLETYGPPSVTLISC 904
Db 853 NSDVERKSCSTQALVFTASTIAYDLPSTGASPAVPTESCHLSVAFVPC 906

RESULT 3
JG4301
follicle-stimulating hormone receptor
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 16-Nov-1995 #sequence.revision 08-Feb-1996 #text.change 21-Jan-2000
C/Accession: JG4301
R:Remy, J.-J.; Labib-Manssals, Y.; Yezle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe
Gene 163, 257-261, 1995
A>Title: The porcine follicle-stimulating hormone receptor: cDNA cloning, functional expression and characterization of the receptor
A:Reference number: JG4301; MUID:96011644
A:Accession: JG4301
A:Molecule type: mRNA
A:Residues: 1-694 <REM>
A/Cross-references: GB:L31966
A/Experimental source: ovarian granulosa cells
A/Comment: This receptor belongs to the family of the G-protein coupled receptors. It
encodes a protein that is involved in the regulation of the G-protein coupled receptors. It
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
C/Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <IOB>
F:170-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR>
F:366-388/Domain: transmembrane #status predicted <TM>

```


Query Match	10.4%	Score 525;	DB 2;	Length 694;
Best Local Similarity	23.8%	Pred. NO. 1, De-28;		
Matches	185;	Conservative 122;	Mismatches 305;	Indels 164;
				Gaps 21;

RESULT 6
JC1493
folliotropin receptor - sheep

Query Match	10.3%	Score 524.5;	DB 2,	Length 695;
Best Local Similarity	26.4%	Pred. No. 1.2e-28;		
Matches	175;	Conservative 101;	Mismatches 273;	Indels 113;
				Gaps 18

```

QY 261 ELGHNHNKIKPKAFKNGNLTOTIHFYDNI-QVGRSAQIYLPKHTLTSLINGAMIQ 319
Db 50 ELKRVLTKRKIPBCAFSGFDLEKTIETSONDVLEVIANFVSNPKLEIRIEKANML 109
QY 320 EF-PDLKGTTSLELTTLTRAGIRLPLPSGMCODPLRYLSELSHNOIEELPSLRHCKLEE 378
Db 110 YIDPD-----AFQNLRLRLILISNTGIKLPAVHKIOSLOK 146
QY 379 I--GLQHN-RWELGATFSOLS-SLOALDISWNAIRSIHPEAF-----STL 421
Db 147 VLDDIODNINIHTVERNSFMGLSEFESMTVWLSKNGIOEIHNCAGFNGTODELNDSNWL 206
QY 422 HSL-----VKLDIDNOLTTPLPLAGGIMHLKLGNNLALSQAFSKDSPPKRI 470
Db 207 EELPNVFOGASGPVILDISRTIRSLPSGLEMLKIRAKSTYHLKPLPSLEKVTTLVE 266
QY 471 LEVPYVQCCPYGMCASFRRASGQWEADLH-----LDDEESSKRPLGLARQA 519
Db 267 ASLTYPHSCAF-----ANMRQTSDLHPICNKSILRQEVDMQARQIRSLAED 318
QY 520 ENNYDDDLDELQLEMEDS--KPHPSVQCSPPRGCFKCEYLFESMGIRLAWAIVLSVL 577
Db 319 EBSYAGFDMTSEFYDLCSEVVDYVTCSEPRAFNCEDEIMGYDILRVLLIMFISILAIT 378
QY 578 CNGVLVLFYFAGGAPRLPVPKFFVYGAJAGANTLIGISGLASDALTLFGQFSEYGAWE 637
Db 379 GNILVAV-ILITSQKYLTVPRFLMCNIAFADLCIGITULLILASVDVHTKSQYHNVAIDMO 437
QY 638 TELGCRATGFLAVLGSSEASVLLITLAAVQ-----CSVSYSC--VRAQKSPSLGSYRA 688
Db 438 TGAGCDMAEFVFEVASELSVYUJLTALTLEBWHHTITHMOJCECKHVHNHNASIMLVGWV-- 495
QY 689 GVLGCLALAGLAAALPLASVGEYGASPLCLPYAPREBQAPALGFTVALVMNSFCFLVVA 748
Db 496 -----FAFAVALPFIETSISSYMKVSIPLM--DIDSPSLQIYVMSLVLMLNVAFAVVIC 546
QY 749 GAYIKYLCDLPRGDF-EAVWDCAMVRHVAVMLPADGLGYCPVAFVAFSASMLGLEPVRPEA 807
Db 547 GCYTHIYLTVRNPNTSSSDTKIARMAMLITTDCLMAPISFFAFASLAKVPLITVSK 606
QY 808 VKSVLLVAVLPACINPLIYLLFNPHFRDRLRLRPRAGDSGFLAYAAAGELKSSCDST 867

```


F.596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match	10.2%	Score 518;	DB 1;	Length 695;
Best Local Similarity	22.2%	Pred. No. 3,6e-28;		
Matches 189;	Conservative 127;	Mismatches 300;	Indels 236;	Gaps 25;
QY 39 CHCEDGIMASACSELSGSAVGGDDLPYAYLDLSNNLTETLOPGLFHHFLFLEEDRLS 98				
Db 23 CHCSNRYFL---CQSKVTEIPSDIPRNAIEKRFVLTAKRVIQKQ----- 64				
QY 99 GNNLSHIPQAFSGLYSL-KILMLQNNLOGGIPAEALWELPSLOSULDA--NLISLVE 155				
Db 65 -----AFSGFDLEKIEISQNDVLEIADVPSNLPKHETIEKANNTLYINPE 114				
QY 156 RSPEGLSLRHMLDNDNALTEIPRALNNLPALQAMTLALNRISSHIPDYAFONILSLVVL 215				
Db 115 -----AFQNLPMNQYLLISNTGIRKHLDP----- 137				
QY 216 HLHNNRIQHLGTHSPEGLHNLLETLDNTNKNLOEFPVIRLIGLOELGFHNNIKAIPEK 275				
Db 138 -VHK-----HSLQKVLID-----IQD-----NINHTERN 163				
QY 276 AEMGNPLIQTHFYDNPPIQVGRSAFOYLPKHTLSLNGAMDIOEPDLKGTSTSEILTYL 335				
Db 164 SFVGS-----FESVILMLNKNQIO--ETHNCAFNQ-----TQDELNL 200				
QY 336 TRAGIRLLPSGMCQDLPRLVLELSHNOIEELPSLRQCKLEIGLQHNRIWEIGDTS 395				
Db 201 S-----DNNNLELPN-----DVH 215				
QY 396 QLSLQALDLSMNAIRSHIPEAFSTLHSLVKLDITDNOITLPLAGLGIMHLKIKGNLA 455				
Db 216 GASGEVILDISRTRIHS-----LPSYGLERLKKLRARSTYN 251				
QY 456 LSAQSFKSDSFKRLRIEYPYAVQCCPYGMCASFKASGQWEAEDH-----LD 504				
Db 252 LKTLPTLEKVALMEASLITPSHCAF-----ANNRQISELHPICNKSILRQEVGY 303				
QY 505 EESSKRPLGLLARAENHYDODLDELQLEMEDSKPH--SVQCSPTPGPFKCEYLFESW 562				
Db 304 MTQTRGORSSLAEDNESSSRGFMTRYTEFPDYDLCNEVADVATCSPKPAFNCEDIMGN 363				
QY 563 GIRLAVALVILSVLCNGVLVLFVAGGAPAPLPVKPFVGAAGANTLTGISCGLASVD 622				
Db 364 ILRLVIMFISLATIGN-IIVLVILITTSQYKLYIPRLIMCMILAFADLCIGLYLLIASVD 422				
QY 623 ALTFGQSESEGARRETGLCRATGFLAVLGSSEASVLLTLTAAYQ-----CSVSVCYRA 676				
Db 423 IHRKSQYHNVAIDMOTAGCDMAAGFTVFASFELSYVTLTATLIERWHITHAMQDLC--- 479				
QY 677 YGRSPSLGSVRAGVIGCLALAGLAALPLIASVGYGASPLCLPAPRPGQPAALGFTYAL 736				
Db 480 --KVQLRHAASVWVMGI-FAFAALPFIIGISSYMKVSTCLPW--DIDSPLSOLYWSL 534				
QY 737 VMANSFCFLVACAYIKLYCDLPRGDF-EAVWDCAMVRHVAMLTFADGLYICPAFLSFA 795				
Db 535 LVYLVLAFLVAGCYIHIYLVLRNPNIVSSSSDTRIARRAMLLFTFDLCMAPISPFAIS 594				
QY 796 SMUGLPEVTPPEAVKSVLYLVLPACLNPLLYLLENFHEFDLRLRPRRAGDSPLAYAA 855				
Db 595 ASLKVPLIYVSKAKILLVLFHPIINSCANPPLIYIFTKNRRDFPILSKG---CYEM 649				
QY 856 AGELEKSSCDST 867				
Db 650 QAOIYRTETTSR 661				

RESULT 9

foliitropin receptor precursor - crab-eating macaque

N/Alternative names: follicle-stimulating hormone receptor (FSHR)

[Species: Macaca fascicularis (crab-eating macaque)]

[Date: 10-Sep-1999 #sequence revision 10-Sep-1998 #over abnno 10-Sep-1998

C:Accession: JN0898; S36452
 R:Gromoll, U.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.
 Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993
 A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of th
 A:Reference number: JN0898; M01D:94071854
 A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CA52463.1; PID:9396802
 A:Note: the authors translated the codon AGT for residue 488 as Arg
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follicleotropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein;
 F:1-17/Domain: signal sequence #status predicted <Sig>
 F:18-695/Product: follicleotropin receptor #status predicted <PRH>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
 F:167-387/Domain: transmembrane #status predicted <TM1>
 F:399-421/Domain: transmembrane #status predicted <TM2>
 F:444-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191-199,293,318/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:556/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 10.2%; Score 515; DB 1; Length 695;
 Best Local Similarity 22.7%; Pred. No. 5, 8e-28;
 Matches 195; Conservative 117; Mismatches 298; Indels 248; Gaps 26;

39 CHCEDGIMSDCESELGSAVPGDLPITAYLDSMNNLTLEQLFHLFLERLS 98
 23 CHCSNRFL--QESKVEIRPSDL-PRNAI-----ELRV 54
 99 GNHLSHIPGAFSGLYSL-KILMIONNOGIGIAEALMELPSLOSRLDA--NLISLVE 155
 55 HTKLEIVIGKGAISGFGLEKIEISQNDVLEIVDFSNIPKHEIRIEKANLLTINDE 114
 156 RSFELSRLHMLDNLATFIPRALNLPALQAMTLALNRIISHIPDAFQMLTSLVYL 215
 115 -----AFQNPENLRYLLISMTGKIHLPD----- 137
 216 HLHNNRIQHLGTHSFEGLHNETLDLNTKQEPFAVIRLGLQELGFHNNIKAIPEK 275
 138 -VHK-----LHSFOKY-----LDDIQ-----DNINHTERN 163
 276 APMGNPLQTHFYDNPLOFVRSAPQYLPKHLTSLGANDIOEPDLKGTSLLETLT 335
 164 SFEGLS-----FESVILMLKNGIQ-----ELHNCAPNG-----TQDELNTL 200
 336 TRAGIRLPSGMCQQLPRLVLELSHNQIEELPSLHRQKLEIGLQHNRIWEIGADTFS 395
 201 S-----DNNNLEELPN-----DVFH 215
 396 QUSLQALDLSMNAIRSIHPEAFSTLSLVKLDLTNDQTLPLAGLGMLHKLKGLA 455
 216 GASGPVLDLSKTRIRHS-----PSSYLENKLKLRARSTYN 251
 456 LSOAFSKDSFPKRLILEVPAVAYOCPCYGMCAAFKASGOWEAEEDLT-----DD 504
 252 LKRIPLSEKLVALMEASITVPSHCAP-----ANMKRQISLHPIKCNKSLROEVDY 303
 505 EESSKRLGLILARQENHYDQDDLEQLMEDESKRPH--SVCCSPTPPGFPCCEVLFESM 562
 304 MTQTRGQRSLAEDNESSYGRGDMTYAEFDYDLCEVVDVTCSPKPDANFPCEDILGYN 363

QY 563 GIRLAWAIVLVLGNGVLTTFEAGGAPLPVPYKVVVAGIAGANTLTGISCGLASVD 622
 DB 364 ILKVLIMFISILATITNTIIVLVLT--TTSQYKLTVPRLMNCNLAFAFDICIGYLLIASVD 422
 QY 623 ALTFGESEFGANWETGLCRATGFLAVIGSEASVLLTLAAVO-----CSVS 670
 DB 423 IHRKSQHNNAIDMOTGAGDAGFEFTVASELSVYTLTAITLEKWHITTAAMQDCVH 482
 QY 671 VSCVRAVYKSPISGSRVAGVIGLALAGLAALPLASVEGASPLCLPYAPPEGOPAL 730
 DB 483 VR-----HAASVMWGMV--FAFMAALPIFGISSYMKVSIPLM--DIDPSLSQ 528
 QY 731 GFTVALVMNSFCEPLVAGAYIKLYCDLPGRDF--EAVWDCAMVRHAMLIFADGILYCPV 789
 DB 529 LYMSLVLVNLAVLVIGCGTHTLYVRNPNTVSSSDTRIKRAMMLTFDFELCMAP 588
 QY 790 AFLFSAMGLFPTVPEAVKSVLLVLPDAGCNPLLYLTFNHFDDLRRLPRAGDSG 849
 DB 589 SFFAISASLKVPLITVSRAKILLVLFYIPNSCANPFLYAFITNFRDPTLLSKFG-- 645
 QY 850 PLAYAAAGLEKSSCDST 867
 DB 646 --CYEMQAOIYTERTSST 661

RESULT 10.

Thyrotropin receptor precursor - dog
 N:Alternative names: thyroid-stimulating hormone receptor; TSH receptor
 C:Species: Canis lupus familiaris (dog)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
 C:Accession: J40077; S06933
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van S
 Science 246, 1620-1622, 1989
 A:Title: Molecular cloning of the thyrotropin receptor.
 A:Reference number: J40077; M01D:90084524
 A:Accession: J40077
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAR>
 A:Cross-references: GB:X29957, NID:9164098, PIDN:AAA30901.1; PID:9164099
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; van S
 Nucleic Acids Res. 17, 10493, 1989
 A:Title: Nucleotide sequence of the dog thyrotropin receptor cDNA.
 A:Reference number: S06933; M01D:90098886
 A:Accession: S06933
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAR>
 A:Cross-references: EMBL:X17146; NID:98449; PIDN:CA53026.1; PID:9850
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-764/Product: thyrotropin receptor #status predicted <MAT>
 F:71-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:201-225/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

Query Match 9.7%; Score 490.5; DB 2; Length 764;
 Best Local Similarity 22.7%; Pred. No. 3, 4e-26;
 Matches 211; Conservative 131; Mismatches 353; Indels 233; Gaps 30;

4 PPGRLALMLCAALCASRRAGAPQPGPTACPAQC-QEGLIMLSADCELSGASVPG 62
 DB 3 PPLHLALLALPRLSIGKGCSP-----PECHQEDDFRT--CKDI----- 44
 QY 63 DLDELAVLDSMNNLTLEQLFHLFLERLSGNHLSHIPGAFSGLYSIKILMQ 122
 DB 45 -----HRIPLPST-QTLKIE-----TQKTPSAFENLPIISRIVLS 84


```

QY 123 NNQIGIPAEALMELPISQSLRLDANLISIVPERSEEGLSIRHML-DNALTEIPYRA 181
Db 85 -----IDATLORL-ESHSEYMLSKMTHIEIKNTSLSDIPA 121
QY 182 LNNPALQAMTLALNRISHIDY-----ARQNL-TSLVY 214
Db 122 LKEPLKFLGIDFTMGGLVFPDVKVYSTVDFILEITDNPYASIPANNAQICNLTIT 181
QY 215 LHLNNRIQHLGTHSEFEGHNLFTLDNLYNKLOEFPVAIRLGLQLGFGHNNNIKAPE 274
Db 182 LKLTNNGFTSIQGHAFNGT-KLDVYLNKNKY-----LSAIDK 218
QY 275 KAFMGNLQTIHFYDNP-----IOFVGRSAF-OYLPEKLTLSLNGAMDOIEPDLKGT 328
Db 219 DARG-----VYSGPTLLDVYSTSVTALPSKGLHKLKELIARNTWLKLP-----L 265
QY 329 SLEILTLFRAGIRLPSGMCQQLPRLRVLELSHNOIEELPSLRCKQLKEIGLQHNRIWE 388
Db 266 SLSEFLHLTRADLS-YPSHCC-----AFKNOKKIRGI-----295
QY 389 IGADTFQSLSQALDLSMNAIRSHIP-EAFSTLSLVKLDITDNLQTLPLAGIGLHM 447
Db 296 -----LESIMCNESISIRSLRQKSVNTLNGPFDQYEYEL-----GDSH 334
QY 448 LKLGKGLALSOAFSKDSFPEKRLILEVPYAYQCCPYGMCASFPGASQWAEEDLHDEES 507
Db 335 AGYKDN---SOPQDDNSHHYV-----FPE-----EOEDEILFGQE 369
QY 508 SKRPLGLARQAEHNYDODLELOEMEDSKRHPSPVOCSPFPKPEYLFESGIRLA 567
Db 370 LKNDPEETLOAFDSHYDYTV-----CGNEDMCTPRSDENPCEDIMGKFKELIV 420
QY 568 WVAIVLLSVLNGVLLVTFVAGGAPARLPVKVYGAINGANTLIGISGLASVDALTFG 627
Db 421 VWFVSLALIGNVFLIYLLSHKLYP-RFLMCNLAFADFCMGYLLTLLASVDLYTHS 479
QY 628 QFSEKGAWEIGLCRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAVGSLSGSVR 687
Db 480 EYVHAIADWQPGPCNTAGFEYVASELSYTLTITLERMYATTFMRDRLKRLHAY 539
QY 688 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPREGOPALAGFTVALMNSFCFLVY 747
Db 540 AINMGWVCCFLALPLRVGISSTAKVSTCLPM-DTEPLALAYITLVLLNVAFLIV 597
QY 748 AGAVIKYICLPKQDEE-AVWDCAMVRHVAWLIFADGILYCPVAPISFASMLGLFPVYFE 806
Db 598 CSCYVAKYITVRNPOYNPGCDKTRIAKRMAYLFTDFMCMAPISFAYLSLMMKPLITVY 657
QY 807 AVKSVLLVLPPLACNPLIYLLFNHFRDRLRLPRAGDSGLAYAAAG-----857
Db 658 NSKILVLPPLNSCANPLIYLFETKAFQRDVFLISKFGICKRQAOAVRGQVSPKNSA 717
QY 858 --ELEKSCDSQOALVAFSDVDLILEAS 883
Db 718 GIOIQKVTDRMROSLPNNODEYELLENS 745

RESULT 11
148882
thyrotropin receptor precursor - mouse
N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor
C:Species: Mus musculus (house mouse)
C:Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 13-Aug-1999
R:Stein, S.A.; Oates, E.L.; Hall, C.R.; Grumbles, R.M.; Fernandez, L.M.; Taylor, N.A.; F
Mol. Endocrinol. 8, 129-136, 1994
A:Title: Identification of a point mutation in the thyrotropin receptor of the hyl/hyt H
A:Reference number: A54271; MUID:94224232
A:Accession: 148882
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-764 <RES>

```

```

A:Cross-references: EMBL:U02602; NID:g575923; PIDN:AB60455.1; PID:g575924
C:Genetics:
A:Gene: TSHR
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C:Keywords: G protein-coupled receptor; transmembrane protein
F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F:201-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

Query Match          9 6%: Score 486; DB 2; Length 764;
Best Local Similarity 21.7%: Pred. No. 7.1e-26;
Matches 211; Conservative 132; Mismatches 346; Indels 284; Gaps 30;

QY 4 PGLRALMCAALCASRAGAPQPGPTACPAQCHC-OEDGIMLSADCELGSLAVPG 62
Db 3 PGLLILVLLALSRSLR-----CKEASPCCECHQDDFRVY--CKEL-----44
QY 63 DLDDLAVLDSMNNITFLOGLFHHLRLEELRLSGNHSHTPGQAFSGLYSLKILMQ 122
Db 45 -----HRPSLPS-----TQTLKLTHTKTPISLAFSSLSLRISRYLS 84
QY 123 NNQIGIPAEALMELPISQSLRLDANLISIVPERSEEGLSIRHML-----WLDNA 173
Db 85 -----IDATLORLEP-HSFYNLSKMHIEIKNTSLTYIDPDA 121
QY 174 LTEIP-----VRANNPALQAMTLAL-----NRISHIPDYAFQNL-TSLVY 214
Db 122 LTELPLKFLGIDFTMGGLVFPDVKVYSTVDFILEITDNPYASIPANNAQICNLTIT 181
QY 215 LHLNNRIQHLGTHSEFEGHNLFTLDNLYNKLOEFPVAIRLGLQLGFGHNNNIKAPE 274
Db 182 LKLTNNGFTSIQGHAFNGT-KLDVYLNKNKY-----LTAIDN 218
QY 275 KAFMGNLQTIHFYDNP--IOFVGRSAFOYLRK-----LHTLSLNGAMDOIEPDLKGT 328
Db 219 DARG-----VYSGPTLLDVYSTSVTALPSKGLHKLKELIANDWTWLKLP-----L 265
QY 329 SLEILTLFRAGIRLPSGMCQQLPRLRVLELSHNOIEELPSLRCKQLKEIGLQHNRIWE 388
Db 266 SLSEFLHLTRADLS-YPSHCCAFKNOKKIRGILIESLMCNESIRILRQKSVNIIIRGPIYQ 324
QY 389 -----IGADTFQSLSQALDLSMNAIRSHIPAFSTLSLVKLDITDNLQTLPL 438
Db 325 EYEDPQGNISVG--YKONSKFQ-----ESPNSHHYV-----354
QY 439 LAGLGLMHLKGLMALSQAFSKDSFPEKRLILEVPYAYQCCPYGMCASFPGASQWAE 498
Db 355 -----FPE-----EOE 360
QY 499 DLHLDDESSKRPGLGLARQAEHNYDODLELOEMEDSKRHPSPVOCSPFPKPECEYL 558
Db 361 DEAVYFGQELKNPQDEETLOAFESHYDYTV-----CGDNEDMCTPRSDENPCEDIL 411
QY 559 FESWGIRLAWAIVLISVLCNGVLLVTFVAGGAPARLPVKVYGAINGANTLIGISGL 618
Db 412 MGYEFLRIVWFVSLALIGNVFLIYLLSHKLYP-RFLMCNLAFADFCMGYVLLLI 470
QY 619 ASVDALTFQGFSEYGAWEIGLCRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAAG 678
Db 471 ASVDLTHSEYVHAIADWQPGPCNTAGFEYVASELSYTLTITLERMYATTFMRDRL 530
QY 679 KSPSLSVRAGVYGLCLAGLAAALPLASVGEYASPLCLPYAPREGOPALAGFTVALVM 738
Db 531 RKIRLRLHATTIAGGAVSCFLLALPVGIISSYARVSTCLPM--DTDPPLALAYIVLVL 588
QY 739 MNSFCFLVVAAYIKLYCDL-----PRGDFEAVWDCAMVRHVAWLIFADGILYCPVAP 792
Db 589 LNVVAVFVVVVCYKVIYITVRNPOYNPRDK-----DTIARMAVLIFTDFMCMAPISFY 643

```



```

QY 491 ASGQEAEDHLHDE---ESSKRPLGLLARGAENHYDODLEDESKPHBSVQCS 546
  |||
Db 402 -GGQ-----PDDVGFGETLKNPEDETSQDFSDRYDVCE---EGBE-----VICA 444
QY 547 PPGPFKECEYLFESMGIRLAWAIVLISVLCNGVLTVFAGGAPLPDPVFFVNGALING 606
  |||
Db 445 PAPDEFNCEDEMGFSFLEKSVWFEVSLAAVGN-MVALLVLTLSHYKLSVSRFLMCHLAF 503
QY 607 ANTLTGISGLASVDALTFQFSEYGARWETGLCRATGFLAVGSEASVLLTLTAAQ 666
  |||
Db 504 ADLCMGIVYLLIASVDLTHRAEYVNHAIIDMONGPCGGLAGFTTVAASELSVTLVITILE 563
QY 667 CSVSVCRAIGKS---PSLGSVR-AGVLGCLALAGLAALPLVASVEYGASPLCLPYAP 722
  |||
Db 564 RMYATFAMRLDKRLRPHAAAVMLAGMLFCL---LLAMLEPLVGVSSYOKYSICLPMY 618
QY 723 PEGOPALGFTVAVLMNMSFCLVAVAGAYIKLYCDLPRGDPEA-VWDCAMRHVAMLIFA 781
  |||
Db 619 -DQSTVAQVYIVSVILNITLAFVLCACYIKITVYVNHPRHSGSKDTNIAKRMVLLFT 677
QY 782 DGLLYCPAFLSFASMLGLFPVTPRAVSVLVVLPACILPLLYLFNPHFDRLRL 841
  |||
Db 678 DFLCMAPISFYAAMSAVLDRLPITVNSKILLVFPYPLNSCANPFLYALFTKAFRGDVFIL 737
QY 842 RPRAG 846
Db 738 LSKVG 742

```

```

RESULT 14
177463
luteinizing hormone/chorionic gonadotropin receptor - rat
C/Species: Rattus sp. (rat)
C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C/Accession: I77463
R/Accession: J.T.; Plettila, E.M.; Lakkakorpi, J.T.; Rajanemi, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A>Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an
A:Reference number: I57668; M01D:92347604
A:Accession: I77463
A>Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-700 <RSS>
A:Cross-references: GB:S40803; NID:g252163; PIDN:AAB22680.1; PID:g252164
C:Genetics:
A:Intons: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
F;I53-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```

```

Query Match          9.4%: Score 474; DB 2: Length 700;
Best local similarity 21.5%: Pred. No. 4.3e-25;
Matches 203; Conservative 130; Mismatches 322; Indels 288; Gaps 30;

QY 5 PGLRALMLCAL-----CASRRAGAPQPGPTACAPCHQOEDGIMLSADCSIELGLS 58
  |||
Db 6 PALRQLLVLAIVLLKPSQLQSRELGS-----SRCEPCDCAPDQAL----- 46
QY 59 AVPGDDPLTAVYDLDSNNNTLEIPLGLFHLRLLELRLSGNHLSHIPGAFSGLYS-LK 117
  |||
Db 47 RCPG---PRAGLARLSITLYP-----VKVLPSSQAFRLGIMENVK 81
QY 118 ILMQNNNOGLGIPAEALMELPSLOSRLD--ANLISLVRSPFEGSLRHLWMDNAL 175
  |||
Db 82 IELSDGSLERIANAFNDILNLSSELLIQTAKMLVIER-GATTNLPRLKLYLSICNGIR 140
QY 176 ELP-VALNMLPALQAMTALN-RISHIDYAFQNTLS-LVVLHLNRRIOHGLTSHFEG 232
  |||
Db 141 TLPDYTKISSSEFNILEICDNHITTPGNARQGMNNEVTLKLVNGEVEVQSHAFNG 200
QY 233 LNHLELDLNYNKLQEPFAIRLGLRLQELGFMNNIKALPEKAFMGNPLLOTIHFYDNP 292
Db 201 T-TLISLELKEN----- 211

```

```

QY 293 IQFVGRSAFOYLPKHTLSLNGAMDIOEPFDLKGTTSLILITRAGITLLPSGMOQLP 352
  |||
Db 212 -----TYLEKMSGAFQATG-----PS----- 229
QY 353 RLVLELSHNOIELELPSLRCKLEIEIGLQHNRIWEIGADTFESQSLQALDLSNNAIRS 412
  |||
Db 230 ---ILDISTKQLALPS-----HG-----LESIOTL----- 252
QY 413 IHPFAFSTHSLVKLDLDNQLTTLPLAGLGLIMHLKLNALSOAFESKDSFPLRIILE 472
  |||
Db 253 -----TALSYSILKTLT-----SKERFTSLVAT 276
QY 473 VPYAYOCCPYGCASFPEFKASGOWEADHLDESSKRPLGLLARGAENH--YDQDLDEL 530
  |||
Db 277 LTFPHCCAFRLRPK-----KEQNSFSIFENFSKCESTVRKADETLYSAIFEN 328
QY 531 QLEMED-----SKPHSVQCSPTPGPFRCXYLFESMGIRLAWAIVLISVLCNGVLTLT 585
  |||
Db 329 ELSGWDYDYGFCSP-KTLOCAPEPDAPNCEIDIMGYAFRLVILNITLAIKGN-LTVLF 386
QY 586 VFAGGPAFLPVPYKRVVGAAGANTLTGISCGLASVDALTFQFSEYGARWETGLCRAT 645
  |||
Db 387 VLLISRYKLVPFRFLMNLSEFADFCMGLYLLIASVDQTKQYVNHAIIDMOTGSGCGAA 446
QY 646 GFLAVLGEASVYLLTLTAAVOCVSVCRAVAGKSPSISGVAVGLGCLALAGLAALPL 705
  |||
Db 447 GFETTFASELSVYTLTTLTERMHTTYTAVOLDOQLRRLHAIPLMGLMSTLITATMPL 506
QY 706 ASVGEYGASPLCLPYAPPEGOPALGFTVAVLMNMSFCLVAVAGAYIKLYCDLPRGDPEA 765
  |||
Db 507 VGISNYMKVSYICLPM--DVESSTLSQVYILSTILNVAFAVYIACACIYITFAVQNPBELTA 564
QY 766 V-WPCAMVRHVAWLIFADGLILXCPVAFSFSMSMLGLFPVTEAKSVLVVLPACILNP 824
  |||
Db 565 PNKDTKAKKKAALLFTDFTCMAPISFEAIAAEKVPRLITYTNSITLLVLEFPVNSCANP 624
QY 825 LLYLLEFNPHFDRLRLRPRAAGSGPLAVAAAGLEKSSCSTQALVAFSDVDILEASE 884
  |||
Db 625 FLVAIFTRAFGRDPLLLSRFG-----CCKRRA-----E 653
QY 885 AGRPGLETY-----GRP-----SVTLISCOQPGAPR 911
Db 654 LYRKKEFSAYTNSCKNGPFGASKPSQATLKLSTVHCQPIPPR 696

```

```

RESULT 15
A49744
luteinizing hormone/chorionic gonadotropin precursor - rat
N/Alternate names: luteinizing hormone-choriogonadotropin receptor
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-May-1994 #sequence_revision 13-Sep-1998 #text_change 13-Aug-1999
C/Accession: A49744; A40545; A41343; A61453; A32460
R/KO: Y.B.; Ji, I.; Slaughter, R.G.; Ji, T.H.
Endocrinology 128, 2297-2308, 1991
A>Title: Structure of the luteinizing hormone receptor gene and multiple exons of the
A:Reference number: A49744; M01D:91209270
A:Accession: A49744
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-700 <KOO>
A:Cross-references: GB:M68917
A:Note: authors translated the codon CAA for residue 307 as Glu, AAC for residue 355
R/etal-Morris, C.H.; Buczko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.
J. Biol. Chem. 266, 11355-11359, 1991
A>Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.
A:Accession: A40545
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <TSA>
A:Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923;
R/McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosembilit, N.; Nikolic
Science 245, 494-499, 1989

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52; Search time 15 Seconds
(without alignments)
2496.120 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MESPPELRALMLCALCASR.....GGLSGGCGPQPSGLAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.5	10.5	692	1	FSHR_MOUSE
2	526.5	10.4	695	1	FSHR_BOVIN
3	525	10.4	694	1	FSHR_HORSE
4	524.5	10.3	695	1	FSHR_SHEEP
5	523	10.3	695	1	FSHR_PIG
6	518.5	10.2	692	1	FSHR_RAT
7	518	10.2	695	1	FSHR_HUMAN
8	515	10.2	695	1	FSHR_MACACA
9	504.5	10.0	687	1	FSHR_EQUUS
10	490.5	9.7	764	1	FSHR_CANINA
11	486	9.6	764	1	FSHR_MOUSE
12	479	9.5	763	1	FSHR_BOVIN
13	474	9.4	760	1	FSHR_RAT
14	473.5	9.3	764	1	FSHR_SHEEP
15	472	9.3	700	1	FSHR_MOUSE
16	469	9.3	701	1	FSHR_BOVIN
17	467	9.2	696	1	FSHR_PIG
18	466	9.2	693	1	FSHR_CHICK
19	464.5	9.2	764	1	FSHR_HUMAN
20	462.5	9.1	764	1	FSHR_RAT
21	455.5	9.0	603	1	ALS_RAT
22	444.5	8.8	676	1	FSHR_CANINA
23	439.5	8.7	699	1	FSHR_HUMAN
24	439	8.7	605	1	ALS_PAPHA
25	438.5	8.7	603	1	ALS_MOUSE
26	437.5	8.6	605	1	ALS_HUMAN
27	422.5	8.3	567	1	GPV_RAT
28	411	8.1	925	1	GLHR_ANGEL
29	397	7.8	1480	1	SLIT_DROME
30	394	7.8	538	1	FSHR_SHEEP
31	375	7.4	560	1	GPV_HUMAN
32	368	7.3	567	1	GPV_MOUSE
33	364.5	7.2	1315	1	CHAO_DROME

34	363	7.2	536	1	CBP8_HUMAN	P22792	homo sapien
35	363	7.2	662	1	GAPR_HUMAN	Q14392	homo sapien
36	357	7.0	366	1	LSHR_CHICK	Q90674	gallus gall
37	349	6.9	713	1	GACL_HUMAN	Q75325	homo sapien
38	348.5	6.9	1115	1	GPCR_TYWT	P46023	lymaea sta
39	348	6.9	782	1	CHAO_TRICA	P82963	tribolium c
40	346	6.8	1049	1	TLR7_HUMAN	Q99YK1	homo sapien
41	340.5	6.7	1039	1	YR71_CABEL	Q09564	caenorhabdi
42	337.5	6.7	582	1	SHO2_HUMAN	Q9UGJ3	homo sapien
43	326.5	6.4	1050	1	TLR7_MOUSE	P58681	mus musc
44	326	6.4	582	1	SHO2_MOUSE	O88520	mus musc
45	324.5	6.4	905	1	TLR3_MOUSE	Q99MD1	mus musc

ALIGNMENTS

RESULT 1
FSHR_MOUSE STANDARD; PRT; 692 AA.
AC P35378: O9QW8: Q9DAC2:
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictropin receptor).
GN FSHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Testis;
RA Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;
RT "Molecular cloning of the mouse follicle stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T transition in exon 7 of the coding sequence."
RT submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schiml L.M., Staudt J., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P., Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
[3]
RN SEQUENCE OF 1-51 FROM N.A.
RP MEDLINE=93093308; PubMed=1459341;
RX Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;
RT "The murine luteinizing hormone and follicle-stimulating hormone receptor genes: transcription initiation sites, putative promoter sequences and promoter activity."
RL Mol. Cell. Endocrinol. 88:55-66(1992).

CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/SH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF095642; AAC67559.1; -
 DR EMBL: AK016635; BAB30351.1; -
 DR EMBL: S49632; AAB24401.1; -
 DR EMBL: M87570; AAB37641.1; -
 DR GCDDB: GCR_0304; -
 DR MGD: MGI:95583; Fshr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS000237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECPT_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 DR Phosphorylation; Repeat; Leucine-rich repeat.
 KW SIGNAL 1
 FT CHAIN 1 692
 FT DOMAIN 18 365
 FT TRANSSEM 366 385
 FT DOMAIN 387 397
 FT TRANSSEM 398 420
 FT DOMAIN 421 442
 FT TRANSSEM 443 464
 FT DOMAIN 465 484
 FT TRANSSEM 485 507
 FT DOMAIN 508 527
 FT TRANSSEM 528 549
 FT DOMAIN 550 572
 FT TRANSSEM 573 596
 FT DOMAIN 597 607
 FT TRANSSEM 608 629
 FT DOMAIN 630 692
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 441 516
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CONFLICT 436 436
 SQ SEQUENCE 692 AA; 77769 MW; 4B5722918053A44 CRC64;
 Query Match 10.5%; Score 532.5; DB 1; Length 692;
 Best Local Similarity 23.7%; Pred. No. 4.1e-29;
 Matches 202; Conservative 124; Mismatches 293; Indels 235; Gaps 30;
 QY 115 STKIMQNNQGGIRPFAIMELP-SIQSLRLDNLISLVERSEGLSSLRHMLDNDNA 173
 DB 26 SNKVFCLQDSKVEIIP---DLPRNAIELEFVLTKLRVLPKSGFSGFGDLEKIEISOND 81
 QY 174 LIE-IPVRLNINLPALQAMTL-ALNRISHIPDVAFOUNLTSVLVHLHNRRIGHLGHSHFE 231

DB 82 VLEVEAEVFNLPRLHRIERIKANLILYNEAONLPSLYLSTNGIKHL-----P 136
 QY 232 GLHNLLETDLNKNKIQEPVAIRTLGLOELGFHNHNNKAIPKAFMGNPLDTTHFYDN 291
 DB 137 AFHRIOSIQ-----KVLIDIOD-----NINIHIAIRNFMGIS-----FES 172
 QY 292 PLOEVGRSAFQYLPRLHTLSNGAMDIOEPFLDKTSLTILTLRAGIRLLPSGCGQL 351
 DB 173 VILNKNKGIO---EINHCARNG-----TOLDENLNS----- 201
 QY 352 PRLVLELSHNOIELPESLARCQKLEIGLOHNRIMEIGADFFSOLSLQALDSMNAIR 411
 DB 202 -----DNNLEBLPD-----DVFQASGVVVIDISRTKY 231
 QY 412 SIHPEASTLSLYKLDLTDNQLTLPILAGIGLMLKLGMLALSOAFKDSFPLRTL 471
 DB 232 S-----LPNHGLENLKLLRARSYYRLKPLSLDKFVMLLEA 267
 QY 472 EVPYAYOCPPYCMGASFPEKASGOMEAEIDLHDDDESSKRPGLLARQENHYDQDEL- 530
 DB 268 SLTTPSHCCAF-----ANWRKOTSELHPTCNKSISR-----QDIDMT 305
 QY 531 -----QLEMEDSKP-----HPSVQCSPTPGPFPCETLFEFSGIR 565
 DB 306 OPDQRYSLVDDEPSYSGSDMLYSEFDYDLCNEFVDYTCSPKPAFPCEDIMKYNILR 365
 QY 566 LAWAIVLISLVCGLVLLVYFAGGPAPLPYKRYVGAIAQANTLTLGICGLLASVDALT 625
 DB 366 VLIMFISILAITGNTVLY-VLTTSQYKLVPRFLMCLMADICIGYLLILASVDIHT 424
 QY 626 FGFSESGARWETGLGRATGFLAVLSEASVLLITLAAY-----QCSYSV 671
 DB 425 KSDYHNVAIDMQTAGGDAGFFTFVFASELV--YTLAATILKRNHTTHAMQLECKYQL 482
 QY 672 SCYRAVKSPLSGSVRAGVLCIALAGIAALPLASVEYGASPLCIYAPPEGPAALG 731
 DB 483 -CHAA-----SIM--VLG-WAFAPAAALFPIFGISSYKVKISCLPM--DIDPLSQL 528
 QY 732 FYVALVMNSECFFVAVGATIKYCDLPRGDF-EAVVMDCAVRIVAVMLIFADGLLYCYVA 790
 DB 529 YVAVALLVNLAVVYIGCTHTIYLRNPNVYSSRRTAKRNATLIFDFLCMAITL 588
 QY 791 FLSPASMLGPEYTPPEAVKSVLLVPLPACLPNPLVLLFNPFRDDLRLRPAGDSGP 850
 DB 589 FFAISASLKVPLTYSKAKILLVLYPINSANCANPFLVAIFTKNRDRDFVL----- 639
 QY 851 LAFAAGLEKSSGDSIQALVAFSDVLLLEASAGRPGLIYGFPSVLLISCOQPGAP 910
 DB 640 -----MSKFGCYEYQAOIYKTIETSSITHNFSRRNP-----CS--SAP 675
 QY 911 RLEGSCHVEPEGNH 924
 DB 676 RVTNNTVLPVPL-NH 688
 RESULT 2
 FSHR_BOVIN
 ID FSHR_BOVIN STANDARD; PRT; 695 AA.
 AC P35376;
 DT 01-JUN-1994 (rel. 29, Created)
 DT 01-JUN-1994 (rel. 29, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicitropin
 DE receptor).
 GN FSHR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.

CC SHRAIN-HOLSTEIN. TISSUE-Ovary, and Testis;
 RX MEDLINE-95127199; PubMed-782612;
 RA Houde A., Lambert A., Saumande J., Silversides D.W., Lussier J.G.,
 RT "Structure of the bovine follicle-stimulating hormone receptor
 RT complementary DNA and expression in bovine tissues.";
 RL Mol. Reprod. Dev. 39:127-135(1994).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L22319; AAC37324.1; -.
 DR HSSP: P23945; 1XUN.
 DR GCRDB: GCR_0766; -.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_5.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PRO0373; GLYCHROMONER.
 DR PRINTS: PRO1143; FSHRECEPTOR.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECPT_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 KM
 FT SIGNAL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SQ SEQUENCE 695 AA; 78084 MW; 18F9DFFEC046380D CRC64;
 Query Match 10.4%; Score 526.5; DB 1; Length 695;
 Best Local Similarity 26.1%; Pred. No. 1, le-28;
 Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15;
 QY 261 ELGFHNNNTKAIPKAFMGNNPLQTIHFYDNP1-QFVGRSAFQYLPKRLTSLGANDIQ 319
 || :||: || | :| : : : : | |||| : : | : :

Db 50 ELRFVLTGLRVIPKGAFCGFDELEKIEISONDVLEIVANFVSNLPKLEHRIERKANLL 109
 QY 320 EF-PDLKTTTSLLEILTTRAGIRLLPSCMGCOOLPRLVLELSHNOIEELPSLRCKLEE 378
 Db 110 YIDPP-----AFONPLMLRLLISNIGIKHLPANVHKQSIQOK 146
 QY 379 I--GLQHN-RWIEGADTFQSLS-SIALDLISWNAIRSIHPEAF-----STL 421
 Db 147 VLLDQIDINIHITVARNSPMGSLFESMTVWLSKNGIOEIHNCFAFGTOLDLEINLSDNSML 206
 QY 422 HSL-----VKDITDNTQTLTPLAGIGGLMHKIKGNLALSOAFSDSPKIRI 470
 Db 207 BELPNDVFGASGPVILDISRTIRISPSYGLNKLKLRKSTYRIKPLSEKVTIVE 266
 QY 471 LEVPAVOCPPYGMCAASFPGKASGOMEADLH-----LDDESSKRRLGLARQA 519
 Db 267 ASLTYPSSHCAP-----ANWRQTSPLHICCKSLIRQVDDMTQARGRVSLAED 318
 QY 520 ENHYDQDDLEQLEMEDSKPHD-SVQCSPTPGPKPCPELFESWQIRLAVAILVSLV 577
 Db 319 EPSYAKGFQVMYSEPDYDLCNEVDYTCSPEDDAFNPCEIDMGDDILRVLIWFISILATY 378
 QY 578 CNGVLLTVFAGGPAPLPVYKRVGAIAGANTLTGICGLASVVALTFGQSESGARME 637
 Db 379 GNILVLY-ILITSQYKLTVPREFLMCNLAFADLCIGYLLILASVDVHTETHTNATIDMQ 437
 QY 638 TGLGCRATGFLAVLGSSEASVLLITLAOVCSVSCVRAYGKSPISGVRAGVLCGLALA 697
 Db 438 TQAGCDAGFFTFVPASLSVTLTALTLERHITIAMOLECKYQLRHAASIMLVGWIFA 497
 QY 698 GLAALPLASVGEASGPLCLPYAPPEGPPALGFTVALVMANSFCELVAVAGATLYCYD 757
 Db 498 FAVALEPPIGISSYMKVSICLPM--DIDSPLSQLYMSLVILVNLVAFVVICGYTHIYLL 555
 QY 758 LRQGF-ELVMDCAVRHVAANLIFADGLIXCPVAFISFPMGLFPVTEAVSVLYLV 816
 Db 556 VRNPNTSSSDTKAKKRAMMLTFDFLCMAPISFAISLAVPLITYSKSKILVLYR 615
 QY 817 PLPACINPLIYLLFNPHFRDRLRLPRAGDSGPLVAAAGELKSSCDST 867
 Db 616 PINSKANPLIYALFTKNFRDFILLSKG-----CYEQAGQYRSETST 661
 RESULT 3
 FSHR HORSE
 ID FSHR HORSE STANDARD; PRT; 694 AA.
 AC P47799;
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
 DE receptor).
 GN FSHR.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NC NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE-94256960; PubMed-8198575;
 RA Robert P., Amselem S., Christophe S., Benifia J.L., Bellet D.,
 RT "Cloning and sequencing of the equine testicular follitropin
 RT receptor.";
 RL Biochem. Biophys. Res. Commun. 201:201-207(1994).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE. AMONG ALL MAMMALIAN FSH RECEPTORS, ON THE HORSE
 CC RECEPTOR DOES NOT BIND LH/CHORIONIC GONADOTROPHIN (CG).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.

291 NPIDPVGKSAQYLPKHTLSLNGAMDQIEFPDLEKGTSTSEILTLFRAGIRLLPSGMCOQ 350

291 NPIQEWGSAFOYLPKIHHTSLNGAMDIQEFPDJKTTSLEILTTRAGIRLLPSGMCQ 350

RT of ovine testicular follicotropin receptor lacking the G protein coupling domains." ; Blochem. Biophys. Res. Commun. 190:888-894 (1993)

RT "Cloning of alternately spliced mRNA transcripts coding for variants
RT of ovine testicular follitropin receptor lacking the G protein
coupling domains.";

RL Biochem. Biophys. Res. Commun. 190:888-894 (1993).
RN [3]

SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RE STRAIN-DORSET-LEICESTER-SUPFOLK 1; TISSUE-Testis;
 RC MEDLINE=98031015; PubMed=9364440;
 RA Varney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.,
 SA Sairam M.R.;
 RT "Molecular cloning, structure, and expression of a testicular
 RT folliculin receptor with selective alteration in the carboxy terminus
 RT that affects signaling function.";
 RT Mol. Reprod. Dev. 48:458-470(1997).
 RL [4]
 RN SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RP STRAIN-DORSET-LEICESTER-SUPFOLK; TISSUE-Ovary;
 RC MEDLINE=20391225; PubMed=10527886;
 RA Babu P.S., Jiang L., Sairam A.M., Touyz R.M., Sairam M.R.;
 RT "Structural features and expression of an alternatively spliced growth
 RT factor type I receptor for folliculin signaling in the developing
 RT ovary.";
 RT Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RL [5]
 RN SEQUENCE OF 1-51 FROM N.A.
 RP MEDLINE=98031017; PubMed=9364442;
 RA Sairam M.R., Subbarayan V.S.R.;
 RT "Characterization of the 5' flanking region and potential control
 RT elements of the ovine folliculin receptor gene.";
 RL Mol. Reprod. Dev. 48:480-487(1997).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of isoform FSH-R1 is mediated by G proteins which activate
 CC adenylylate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
 CC this does not result in activation of adenylylate cyclase. Isoform
 CC FSH-R3 may be involved in calcium signaling. Plasma membrane
 CC (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2,
 CC FSH-R3 and FSH-R4; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
 CC testis, but not in kidney.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSR SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: I07302; AAA31525.1; -
 DR EMBL: L12766; AAA31523.1; -
 DR EMBL: L12767; AAA31524.1; -
 DR EMBL: L36115; AAK70667.1; -
 DR EMBL: A131735; CA10495.1; -
 DR EMBL: AF090438; AAC61749.1; -
 DR PIR: JCI493; JCI493.
 DR HSSP: P23945; JXUN.
 DR GCRDB: GCR_0496; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 5.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00373; GLYCHROMONER.
 DR PRINTS: PR01143; FSHRECEPTOR.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 695
 FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
 FT

TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT VARSPPLIC 126 133
 FT VARSPPLIC 135 695
 FT VARSPPLIC 224 259
 FT VARSPPLIC 260 695
 FT VARSPPLIC 643 670
 FT VARSPPLIC 671 695
 FT VARSPPLIC 695 AA; 78237 MM; FBF5D89D88C0D4B CRC64;
 SQ SEQUENCE
 Query Match 10.38; Score 524.5; DB: 1; Length 695;
 Best Local Similarity 26.48; Pred. No. 1.5e-28;
 Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;
 261 ELGFNNNNKAIPEKAFNGNPLQTHFYDNP1-QFYGRSAFOYLPKLTLSLGANDIQ 319
 50 ELRFVLTIRVPEAGSGFDLEKIEISQNDVLEIVANFSLPKLHEIRIKANNLL 109
 320 EF-PDLKGTSTLEILLTRAGIRLLPSGMCQQLPRLVLEISHNOIEBPSLHCQKLE 378
 110 YIDPD-----AFONLPLRYLLISNGIKHLPVHRKIQSLQK 146
 379 I--GLQHN-RWIEIGADTFESQLS-SLOALDLSMNAIRSIHPEAF-----STL 421
 147 VLLDIDQNTINITHVERNSFMGISFESMTWLSKNGIQEIHNCATFNGTDLDELNSDNL 206
 422 HSL-----VKLDLTDNQLTTLPLAGIGIMHLKGNLALSOAFSKDSFPKRI 470
 207 EELPNDVFGASGPVILDISRTIRSLPSYGENLKLRAKSTYHLKLPSEKRVTLVE 266
 471 LEVPAVOCPPYGMCAFPKASGQWEADLH-----LDDDESSKRLPLGLARQA 519
 267 ASLTYSHCCAF-----ANWRQTSDLHPICNKSILRQEDVDMTQAKGORISLADD 318
 520 ENHYDDDELQLEWEDS--KPHSVQSPTPGPKPCPEYFESWGIIRLVAWAILSVL 577
 319 EPSYAKGDMKSEFYDLCSEVDVTCSPEDAFNPGCDINGYDLRLVIFISILAIT 378
 578 CNGVLLTVFAGGPAPLPVKRYVGAIGANTLTIGISGLASVYDALTPGQFSEGARME 637
 379 GNILVIV-ILITSQYKLTVPRLFMCMIAFADLCIGTYLLILASVDVHRKSQTHNVAIDMQ 437
 638 TGIGCATATGLAVLGEASVLLTLAAVO-----CSVSVSC--VRAYGKSPSLGSVRA 688
 438 TGAGCDAAGFTVFASBLSTYTTTATLTERWHITIAMQLECKVHRAASIMLVGVW-- 495
 689 GVLGIALAGIAAALPLASVGEASGPLCLIPVAPBEGPAAIGFTYALVMMNSFCLVYA 748

```

Db 496 -----FAFAVALPEIFGSSYMKVSIQPM--DIDPSLSQLYVMSLVLNLAFAVVIC 546
Qy 749 GAVIKYLCPLRGDF-EAVYDCAMVRVHVMILADGLICPVAFSLFSPASMLGLFPVTPPEA 807
Db 547 GCETHYLVVRNNNTSSSDRTKIRAKRMAMLTFTDFLCNAPISFFAISLKLKPLVTVSK 606
Qy 808 VASVLVLPACLPNLLYLFENPHRDDLRRLRPACGSLAVYAAAGLEKSSCDST 867
Db 607 SKLLVLFEPYINSCANPFLYAFTRNFRDFFIL-----LSKFGCYEV 649
Qy 868 QA 869
Db 650 QA 651

RESULT 5
FSHR_PIG STANDARD: PRT: 695 AA.
AC P49059; 077514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictropin
  receptor).
GN FSHR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA MEDLINE=96011644; PubMed=7590277;
RA Remy J.J., Labhidi-Mansali Y., Yerde M., Bozon V., Couture L.,
  Pajot E., Grebert D., Salesse R.;
RT "The porcine follicotropin receptor: cDNA cloning, functional
  expression and chromosomal localization of the gene.";
RL Gene 163:257-261(1995).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
  la Barbera A.R.;
RT "Porcine follicle-stimulating hormone receptor.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
  OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
  ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL, L31966; AAA86933.1; -.
CC EMBL, AF023377; AAC24981.1; -.
CC HSSP; P23945; 1XUN.
CC GCRB; GCR_1561; -.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000372; LRR_Nterm.
CC Pfam; PF00001; 7tm_1; 1.
CC Pfam; PF00560; LRR; 4.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00373; GLYCOPROMER.

```

```

DR PRINTS; PR01143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
  Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1..17
FT CHAIN 18..695
FT DOMAIN 18..366
FT TRANSMEM 367..387
FT DOMAIN 388..398
FT TRANSMEM 399..421
FT DOMAIN 422..443
FT TRANSMEM 444..465
FT DOMAIN 466..485
FT TRANSMEM 486..508
FT DOMAIN 509..528
FT TRANSMEM 529..550
FT DOMAIN 551..573
FT TRANSMEM 574..597
FT DOMAIN 598..608
FT TRANSMEM 609..630
FT DOMAIN 631..695
FT REPEAT 44..68
FT REPEAT 69..93
FT REPEAT 119..143
FT REPEAT 170..192
FT REPEAT 193..216
FT REPEAT 218..240
FT DISULFID 442..517
FT CARBOHYD 191..191
FT CARBOHYD 199..199
FT CARBOHYD 293..293
FT CONFLICT 2..2
FT CONFLICT 13..13
FT CONFLICT 60..60
FT CONFLICT 166..166
FT CONFLICT 215..215
FT CONFLICT 247..247
FT CONFLICT 257..257
FT CONFLICT 334..334
FT CONFLICT 349..349
FT CONFLICT 352..352
FT CONFLICT 383..383
FT CONFLICT 407..407
FT CONFLICT 421..421
FT CONFLICT 427..427
FT CONFLICT 435..435
FT CONFLICT 483..483
FT CONFLICT 550..550
FT CONFLICT 586..586
FT CONFLICT 607..607
FT CONFLICT 691..691
SQ SEQUENCE 695 AA; 78172 MW; E9EBED29C79C450 CRC64;

Query Match 10.3%; Score 523; DB 1; Length 695;
Best Local Similarity 26.1%; Pred. No. 1,9e-28;
Matches 171; Conservative 103; Mismatches 254; Indels 126; Gaps 17;

Qy 261 ELGFHNNTKRAPEKAFMGNPILQTHFYDNP- QVYGSAPFYLRKLTLSNGAMDIO 319
Db 50 ELRFVLTKLRLVPKGAFGSGDLEKIEISONDVLEIANVFSLPKLHEIRIEKANMLL 109
Qy 320 EF-PDLKGTSTSEILTLFRAGIRLSPGMCQQLPRVLELSINQTEELPSLRQCKLE 378
Db 110 YIDPD-----AFQNPINRILYLLINSTGVKHLPAVAKIOSLOK 146
Qy 379 I--GLQHN-RIMEIGADTFSQLS-SLQALDSLMAIRSHPEAFS-----TL 421
Db 147 VLLDIQDININIHVERNSFVGLSFESMIILMSKNGIREIHNCAFNGTQDELNLSDNDL 206
Qy 422 HSL-----VKLDITDNLQTLPLAGLGLMHLKLKKNLNLSQLFSDSPFKLRI 470

```

Db 207 EELPNDVFGASGPVILIDISRTIRHSLPSYGLENLKRLAKRSTYNLKKPLSEKFTVLTME 266
 QY 471 LEVPYAVOCCPYKMGASFFKASGOWEADILH-----LDDE 505
 Db 267 ASLTPSHCCAF-----AMMRQISDHPICNKSILRQEVDMTOARGRVSLMEDG 318
 QY 506 ESSKRPLGLLARAENHYDDDLDELQEMEDSKHPSPVCGPPKPCPELFESAGIR 565
 Db 319 ESS-----LAKEDTMYSEDDDLCEVVD-----VISPEDPTNPCEDIGHDILR 366
 QY 566 LAWAIVLTVLNGVLTIVFAGGAPALPVPKVEVGAIGANTLTGISCGLASVDALT 625
 Db 367 VLIWFISILMITGN-ILIVILITISQYKLLPRLMLCNLAFADICIGIYLLILNSVDHT 425
 QY 626 FGFSEYGAEMWGLGCRATGFLAVLGSESVLLTLTA-----VQCSVSYSC 673
 Db 426 KTOYHNAIDMOTGAGCDAGFFVFASELSVYTLTITLERMTITHAMOLOCKV---- 481
 QY 674 VRAVGKSPISGVAVGLGCLALAGLAAALPLASVGEYSGAPLCPYAPREGOPALGFT 733
 Db 482 -----QLRHAASIMLVGMITFAFYALPPIFGISSTMYKSYICLPM--DIDPSQLYV 531
 QY 734 VALVMNSFCELVVAGAYIKLYCDLPGRGD-FAVWDCAMVRHVAMLIJFADGLYCPVAF 792
 Db 532 VSLILVAVLAVFVVICGYTHYLVFRNPNIMSSSDTKIAKRAMLIJFTDPLCMAPISFF 591
 QY 793 SFASMLGLFPVTPRAVAVSVLLVYPLPACLPPLLYLFPNPFRRDLRLRPRAG 846
 Db 592 AISASLKVPLITVSKSKILLVLFYVINSKANPFLYAFITKFRDVFILLSKFG 645
 RESULT 6
 FSHR_RAT
 ID FSHR_RAT STANDARD: PRT: 692 AA.
 AC P20395;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicleprolin receptor).
 GN FSHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Sertoli cells;
 RX MEDLINE=91125358; PubMed=2126341;
 RA Sprengel R., Braun T., Nikolic K., Segaloff D.L., Seeburg P.H.;
 RT "The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA.";
 RT Mol. Endocrinol. 4:525-530(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92149579; PubMed=1738373;
 RA Heckert L.L., Daley I.J., Griswold M.D.;
 RT "Structural organization of the follicle-stimulating hormone receptor gene.";
 RT Mol. Endocrinol. 6:70-80(1992).
 RL MoI. FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: SERTOLI CELLS AND OVARIAN GRANULOSA CELLS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb.sib.ch).
 CC -----
 Db EMBL: L02842; AAA41175.1; -.
 DR PIR: A34548; A34548.
 DR PIR: A41729; A41729.
 DR HSPR: P23945; 1XUN.
 DR GCRDB: GCR_0234; -.
 DR GCRDB: GCR_0456; -.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF000560; LRR_3.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00373; GLYCHROMER.
 DR PRINTS: PR01143; FSHRECEPTOR.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 692
 FT DOMAIN 18 365
 FT TRANSMEM 366 386
 FT DOMAIN 387 397
 FT TRANSMEM 398 420
 FT DOMAIN 421 442
 FT TRANSMEM 443 464
 FT DOMAIN 465 484
 FT TRANSMEM 485 507
 FT DOMAIN 508 527
 FT TRANSMEM 528 549
 FT DOMAIN 550 572
 FT TRANSMEM 573 596
 FT DOMAIN 597 607
 FT TRANSMEM 608 629
 FT DOMAIN 630 692
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFD 441 516
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SO SEQUENCE 692 AA; 77681 MW; 267EA78C7CFD8EC6 CRC64;
 Query Match 10.2%; Score 518.5; DB 1; Length 692;
 Best Local Similarity 23.8%; Pred. No. 3.8e-28;
 Matches 201; Conservative 123; Mismatches 304; Indels 217; Gaps 30;
 QY 115 SIKIMLQNNQGGIPAEIMWELP-SLOSRLDANLISLIPRSPEGLSTLHMLWDNA 173
 Db 26 SNRVFLQDSKTEITPT-----DLPRNAIELRVLTKRVIYIGSGAGGDELEKIEISQND 81
 QY 174 ILE-IPVARNLNPALQAMTL-ALNRISHIDYARQNTLSVYLHLHNNRIQHLGTHSFE 231
 Db 82 VLEVTEADVFSPLPKLHEIRKANNLLYINPEARQNPSTLRYLLISNTGIRKL-----P 136
 QY 222 GLHNLETDLNKNKLOEPVPAIRTKGRLOELCFHNNNKAIPKAPMGNPLOTHFYDN 291
 Db 137 AVHKIOSLO-----KVLIDIOD-----NINHIYARNSEFMGIS-----PES 172
 QY 222 PLOVGRSAFOYLPKRLHTLSNGAMDIQEPDLKGTSLSEILTLRAGIRLPSGMCOOL 351
 Db 173 VILMISKNIE---EIHNCAPNG-----TQDELNTLS----- 201
 QY 352 PRLRVLELSHNOIEELPSLHRCQKLEIGLOHNRIMEIGADTFPSOLSLQALDLSMNAIR 411

```

Db 202 -----DNNMLEEJPN-----DVFQASGSPVILDSRKHV 231
Qy 412 SIHPEAFSTLSLVKDLNDQUTPLAGLGLMLKLGKLNALSGAFESKDSFKRIL 471
Db 232 S-----LPHNGLENLKLKLRARSTYRLKLPMDKRVTLME 267
Qy 472 EWPYAYOCCPYGMCASFRRASGOMEADLH-----LDDEE 506
Db 268 SLTYSHCCAF-----ANLKRQISELHPICNKSILRODIDDMQIGDQRSLIDEP 319
Qy 507 SSKRPLGLLARAENHYDDDELQLEMEDSKHPVSQSPTRPPEPCYEYLESWGIRL 566
Db 320 S---YGGSGDMMYNEFDYLCN---EYVD-----VTCSPKPDANPCEDINGNILRV 366
Qy 567 AVMAIVLLSVLCNGVLYLVFVAGGAPAPLPVKFVGAIGANLITGSCGLASVALTF 626
Db 367 LIMEFSLIATIGNTVLY-VLTTSQYKLTVPRLMCLNLPADLCIGYLLILASVDITK 425
Qy 627 GOFSEYGARWETGLGCRATGFLAVLGESEASVLLTLAAVQ-----CSVSVSCVRAYGS 680
Db 426 SQYHNYAIDMOTGAGCDAAGFFTFVASELSVTLTITLERMHTITHAMQLEC-----KV 480
Qy 681 PSLGVRAGVGLCLAGLAAALPLASVEYCASPLCLPYAPPEGPALGFTVALVMAN 740
Db 481 QLRHAASYAVLG-WTFAPFAALFPIFGISSYKVSICLPM--DIDSPLSQLYVMALVLN 537
Qy 741 SFCEFLVAGAYIKLYCDDLPGRDF-EAWMDCAVRVHVMILFADGLLYCPAFELSPASMG 799
Db 538 VLAFFVIGCGYHILYLVANPTVSSSSTKAKRNATLITFDLMAPISFASISLK 597
Qy 800 LEPVTPAVKSVLLVLPPLACINPLLYLFPNFRDRLRLPRAAGSGPLAYAAAGEL 859
Db 598 VPLIVSRKAKILLVLFYPLINSCANPLFYAIFTKNFRDFEIL-----L 640
Qy 860 EKSSCDSTQALVAFSDVLLILEASRAGRPGLTYFPFVSITLSCQPPAPRLGSHCV 919
Db 641 SKFGYEMQAOIYRT-----ETSSA-----THNF-HARKSHCS-SAPRVTSYVLY 684
Qy 920 PEGNH 924
Db 685 PL-NH 688

```

RESULT 7

FSHR_HUMAN STANDARD; PRT; 695 AA.

AC P3945;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin receptor).

GN FSHR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID:9606;

PN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RX MEDLINE=91222171; PubMed=1709010;

RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;

RT "Cloning and sequencing of human FSH receptor cDNA."

RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=93246012; PubMed=1301382;

RA Keiton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,

RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzaja J.B., Luchette C.A.,

RA Chappel S.C.;

RT "The cloning of the human follicle stimulating hormone receptor and

```

RT its expression in COS-7, CHO, and Y-1 cells.";
RL Mol. Cell. Endocrinol. 89:141-151(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Tilly L.T., Alhara T., Nishimori K., Jai X.-C., Billig H.,
RA Kowalewski K.I., Perlas E.A., Hsueh A.J.;
RL Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-342 FROM N.A.
RC TISSUE=Testis; PubMed=1359889;
RX MEDLINE=93075197; PubMed=1359889;
RA Gromoll J., Gudermann T., Nieschlag E.;
RT "Molecular cloning of a truncated isoform of the human follicle
stimulating hormone receptor.";
RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
RN [5]
RP SEQUENCE OF 1-51 FROM N.A.
RX MEDLINE=95011044; PubMed=7926278;
RA Gromoll J., Dankbar B., Gudermann T.;
RT "Characterization of the 5' flanking region of the human follicle-
stimulating hormone receptor gene.";
RL Mol. Cell. Endocrinol. 102:93-102(1994).
RN [6]
RP 3D-STRUCTURE MODELING OF 49-228.
RX MEDLINE=96363672; PubMed=8747461;
RA Jiang X., Dreano M., Buckler D.R., Cheng S., Yehier A., Wu H.,
RA Handrickson W.A., el Tayar N.;
RT "Structural predictions for the ligand-binding region of glycoprotein
hormone receptors and the nature of hormone-receptor interactions.";
RL Structure 3:1341-1353(1995).
CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: SEROTOL CELLS AND OVARIAN GRANULOSA CELLS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; M65085; AAA52477.1; -
DR EMBL; S59900; AAB26480.1; -
DR EMBL; M95489; AAA52478.1; -
DR EMBL; X68044; CAA48179.1; -
DR EMBL; S73199; AAB32071.1; -
DR PIR; JN0122; JN0122.
DR PIR; JN0122; JN0122.
DR PIR; JN0122; JN0122.
DR GCRDB; GCR 0071; -
DR GCRDB; GCR 0404; -
DR GCRDB; GCR 0588; -
DR GCRDB; GCR 0690; -
DR MIM; 136435; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF000001; 7tm_1; 1.
DR Pfam; PF00560; LRR_4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHROMONER.
DR PRINTS; PR01143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW

```

Db 252 LKKPLETLEKLVALLMSEALVTPSHCCAF-----AMWRQISLHPICKNSILIRQEVDDY 303

QY 505 EESSKRPLGLIARQAEHHYDQIDDELQLEMDSKPH--SYQSGPTGPCKPECEYLESM 562

Db 304 MTQRRGRRSIIAEENNESSYSGKGFDMYTEEDYDLCNEVVDVTCSPKPDARNPCEIDMGYN 363

QY 563 GIRLAWAIVLIVSLQNGVILVTFYAGGPAPLPVKKFVYGAIGAPNLTGISGLIASVD 622

Db 364 ILRLVIMFISLITATGN-IYLVILTTTSQKILYPRFLMCNLAADLCIGIYLLILIASVD 422

QY 623 ALTFGESEYGARREYTGICGRATGFLAVLGSEASVLLITLAAV0-----CSYVSCVRA 676

Db 423 IHTSQYHNVAIDMGTAGCGDAAGFFIYFASLSVYTLTATLERMHTITHAMQLDC--- 479

QY 677 YGKSPSLGSVAVGLVGLCIATLAAGLAALPLASVGEYASPLCLPAAPREGPAAIGFYVAL 736

Db 480 --KVLRRHAASVMWGWG-FAPFAALPFIIFGSISSMKYSIDLP--DIDPLSOLYWSL 534

QY 737 VMANSFCLVYAGAVYIKLYCDLPBGDF-EAYWDCAMVRHVALIFPADGLICPAFLSFA 795

Db 535 LVNLAVLAVVIGCGTIRHYLTVRNPNIVSSSDTRIARKRAMLFTFDPLCAPISFPAIS 594

QY 796 SMLGTFPTVPAVKSVLVLVPLPACINPLDLYLLENFPRDRLRLRPRAGDSGLAYAA 855

Db 595 ASLAKVPLTVTSKAKILLVLFHPINSCANPFLYIAFTKFRNRFDETLISKGS-----CYEM 649

QY 856 AGELEKSSCDST 867

Db 650 QAQIYRTETSSY 661

RESULT 8

ID FSHR_MACFA STANDARD; PRT; 695 AA.

AC P32212;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).

GN FSHR.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=94071854; PubMed=7504463;

RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;

RT "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis.";

RL Biochem. Biophys. Res. Commun. 196:1066-1072(1993).

-I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.

-I- SUBCELLULAR LOCATION: Integral membrane protein.

-I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

FSHR/LSH/TSR SUBFAMILYX

-I- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).

CC

DR EMBL: X74454; CAA52463.1; -.

DR PIR: S36452; S36452.

```

DR PIR; JN0898; JN0898.
DR HSSP; P23945; 1XUN.
DR GCRDb; GCR_0653; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PRO0373; GLYCHOMONER.
DR PRINTS; PRO1143; FSHRECEPTOR.
DR SMART; SMO0013; LRRNT; 1.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSSEM 367 387
FT TRANSSEM 388 398
FT TRANSSEM 399 421
FT TRANSSEM 422 443
FT TRANSSEM 444 465
FT TRANSSEM 466 485
FT TRANSSEM 486 508
FT TRANSSEM 509 528
FT TRANSSEM 529 550
FT TRANSSEM 551 573
FT TRANSSEM 574 597
FT TRANSSEM 598 608
FT TRANSSEM 609 630
FT TRANSSEM 631 695
FT DOMAIN 631 68
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CMC64;

Query Match 10.2%; Score 515; DB 1; Length 695;
Best Local Similarity 22.7%; Pred. No. 6.6e-28;
Matches 195; Conservative 117; Mismatches 298; Indels 248; Gaps 26;

QY 39 CHCOEDGIMLSADCSSEGLSAVPGDLDPLRAYLDLSMNNLTLELQGLFHHLEELRLS 98
DB 23 CHCSNRFL-----COESKATEIPSDL-PRNAI-----ELRFV 54
QY 99 GNHLSHIPGQAFSGLYSL-KILMLQNNOLGIPAPALWELPSLOSLRLDA-NLISLVE 155
DB 55 HTKLRYIQKAFSGFGLLEKIEISQNVLEIADVPSNLPKHEIRLEKANNLLYINPE 114
QY 156 RSEEGSLSLRHLMDNALTEIPYRALNIPALQAMTLALNRISHIDPYAFQNTLSLYVL 215
DB 115 -----AFQNLPLNLYLLISNGIKHLPD----- 137
QY 216 HLHNNRIQHLGTHSFEGLANLETLDLNLYNKLQEPFAIRIGRQELGFHNNNIKAIREK 275
DB 138 -VHK-----HSFOV-----LLDIO-----DINIHITERN 163
QY 276 AFMGNDLQTIHFYDNPIDQVGRSAFQYLPKLTLSINGAMDIOEPDLKGTSLLEITL 335
DB 164 SFVGLS-----FESVILMLNNGIQ-EINNCANG-----TQIDELNL 200
QY 336 TTAGTILPLSGMCQQLPRLRVLEISHNQIEELPSLHRCQKLEIGLOHNRITWEIGADTF 395
DB 201 S-----DNNNLEELPN-----DVFH 215

```

```

QY 396 QLSQIALDLSWMAIRSHHPAFTLSLVLKIDLTNDQTLPLAGGLMLKTKGNIA 455
DB 216 GASPEVILIDISRTIRHS-----LPSYLENLEKILRARSTYN 251
QY 456 LSGAFSKDSFPEKRLLEVPYAVOCCPYGCASFEGKASGOMEAEDH-----LDD 504
DB 252 LKRLPSLEKVALMEASILTPSHCCAF-----ANMRQISELHPICNKSILQOEVDY 303
QY 505 EESSKRLGLIARQAEHNYDDDELQLEMEDSKRP--SVQCSPTPGPKCEYLFRSW 562
DB 304 MTQTRGQRSSLAEDNESSYRGFMTRYAEFDYDLCNEVYDVTCSPKPAFNCEDILGVN 363
QY 563 GIRLAWAIVLISVLCNGVILVTFAGAPRALPVPKFFVGAAGANTLTGSCGLASVD 622
DB 364 ILKRLIFISLITATGNIYLVTL-TTSQYKLVPRFLMCNLAFAADLCIGYLLIASVD 422
QY 623 ALTFGPSEYGARWETGICRATGFLAVLGSSEASVLLITLAAVQ-----CSVS 670
DB 423 IHTKSOYHNVAIDMQTGAGDAAGFFVFASELSVYTLTATLEEMHTITAMQDCKVH 482
QY 671 VSCYRATGKPSLGSVYAGVLCIALAGLAADPLASVGEYGASPLCLYPAPPEGOPAL 730
DB 483 VR-----HAASVMWGW-FAFAALFPFIQISSYMKVSIQCPM--DIDPSLQ 528
QY 731 GFTVALVMNSFCFLVYAGAYIKLYCDLPRGDF-EAVWDCAVHVHVMILRADGLLYCPV 789
DB 529 LYVMSLVLVNLAFAVVICGCTHYLVRNPNIVSSSDRIARAKMAMLIITDFLCMAPI 588
QY 790 AFLSFSMGLGFPVTPPAVKSVLVLPACLNPLLYLTFNPFRRDLRLRAPAGDSG 849
DB 589 SFPAISASLVPVLPITYSKAKILLVFPYNSCAMPFIYATPKNRRDFLLSKFG--- 645
QY 850 PLVYAAAGLEKSSCOST 867
DB 646 --CYEMQAQIYKRTST 661

RESULT 9
FSHR_EOUCAS STANDARD; PRT; 687 AA.
AC 095179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Equus asinus (Donkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9793;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis.
RX MEDLINE=97338913; PubMed=9195473;
RA Richard F., Martinat N., Remy J.-J., Salesse R., Combarrous Y.;
RT "Cloning, sequencing and in vitro functional expression of
RT recombinant donkey follicle-stimulating hormone receptor: a new
RT insight into the binding specificity of gonadotropin receptors.";
RL J. Mol. Endocrinol. 18:193-202(1997).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- FSH/LSH/FSH SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17146; CA35026.1; -
 DR EMBL: X17147; CA35027.1; -
 DR EMBL: M29597; AAA30901.1; -
 DR EMBL: M90047; AAA30902.1; -
 DR PIR: S06933; S06933.
 DR PIR: A40077; A40077.
 DR HSSP: P16473; 1XUM.
 DR GCRDB: GCR_0020; -
 DR GCRDB: GCR_0021; -
 DR GCRDB: GCR_0022; -
 DR GCRDB: GCR_0224; -
 DR InterPro: IPRO00276; GPCR_Rhodopsn.
 DR InterPro: IPRO01611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR PRINTS: PR00373; GLYCHROMONER.
 DR PRINTS: PR01145; TSHRECEPTOR.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 764
 FT DOMAIN 21 413
 FT TRANSSEM 414 441
 FT DOMAIN 442 450
 FT TRANSSEM 451 473
 FT DOMAIN 474 494
 FT TRANSSEM 495 517
 FT DOMAIN 518 537
 FT TRANSSEM 538 560
 FT DOMAIN 561 580
 FT TRANSSEM 581 602
 FT DOMAIN 603 625
 FT TRANSSEM 626 649
 FT DOMAIN 650 660
 FT TRANSSEM 661 682
 FT DOMAIN 683 764
 FT REPEAT 51 74
 FT REPEAT 150 174
 FT REPEAT 176 199
 FT REPEAT 201 223
 FT DISULFID 494 569
 FT CARBOHYD 77 77
 FT CARBOHYD 99 99
 FT CARBOHYD 177 177
 FT CARBOHYD 198 198
 FT CARBOHYD 302 302
 FT VARSPIC 81 105
 SQ SEQUENCE 764 AA; 86483 MW; 49F03B3DBCB65512 CRC64;
 Query Match 9.78; Score 490.5; DB 1; Length 764;
 Best Local Similarity 22.78; Pred. No. 3; 7e-26;
 Matches 211; Conservative 131; Mismatches 353; Indels 233; Gaps 30;
 QY 4 PPGALMICAACARRAGAPGPGPTACAPCHC-OEDGIMISADCESELGLSAVPG 62
 DB 3 PPPLHLALALALPRBSGKGCSP-----PCBQEDDPFRYT--CKDI----- 44
 QY 63 DDDPLAVAYDLNNNTLQPGFLHFLLELRLSGNHLSHIPGAFSGLYSLKILMLQ 122
 DB 45 -----HRIPLPPST-QTLKFI-----TQKTIPLSRAFSNLPNISRIYLS 84

QY 123 NNOLGIPAEALWEIPSLQSLDANLISLVEPRSEGLSSRLHML-DONALTEIPVA 181
 DB 85 -----IDATLQRL-ESHSEYNNLSKMTHEIRNTRSLTSDPDA 121
 QY 182 LNNPALQAMTLALNNISHIPY-----AFQNL-TSLV 214
 DB 122 LKELPLKRLKGLFNGGIGFPYTVKYSTDVFFLEITDNPYMASIPANAFGLCNETLT 181
 QY 215 LHLHNNRIGHLQTSFEGHNLLETLDLNNKLOEFPVALRTGLQELGFHNNIKAIPE 274
 DB 182 LKLYNNGFTSIQGHAFNGH-KIDAVYLNKNKY-----LSAIDX 218
 QY 275 KAFMNPPILQTHFYDNP-----IQFVNSAF--QYLPKHLTSLGAMDIOFEPDLKGT 328
 DB 219 DAFGG-----VSGPFLDVSTYVVALPCKLEHLEKELIARNTWLTKLP-----L 265
 QY 329 SLEILTLTRAGIRLPSGMCQOLPRLRVLELSHNGIEELPSLRCKLEIGLQNNRIWE 388
 DB 266 SLSPHLTRADLS-YPSHCC-----AFKNQKTRGT----- 295
 QY 389 IGADFFSOLSLQALDLSWNAIRSIHP-EAFSTLSLVKLDLTDNQTLPLAGIGLMH 447
 DB 296 -----LESIMCNESIRSLRQKRSVTLNGPPEQEEYVL-----GDSH 334
 QY 448 LKLGKLNALSOAFSKDSPKRLILEVPYAYOCCPYGMCASFKAQGWAEADLHDDDES 507
 DB 335 AGYKDN---SQFQDTDSHSHYV-----FPE-----EQDEILIGFQOE 369
 QY 508 SKRPLGLARQAEHHYDODDELQLEMEDSKPHPSVOCSPTPGPPKCEYLFESGIRLA 567
 DB 370 LKNPEETLQAFDSHYDYV-----CGNEMWCTPSPDENPCEDLMGYFKRLIV 420
 QY 568 VMAIVLTVLNGVLYLVFAGGAPRLPVYFVGALINGANTLIGISGLLASVDALTFG 627
 DB 421 VWEVSLALLNVEFLVLLVLSHKLVLP-RFLMCNLAFCGCMGYLLTLASVDLYTHS 479
 QY 628 QFSEYGAHWETGLCGRATGFLAVLGESEASVLLTLAAVQCSVSVYARVAKSPSLGSR 687
 DB 480 EYVNHAIIDMOWPGPCNTAGFTTFVASELSVYTLVITERRKAITFAMRLDRKIRLRAV 539
 QY 688 AGVIGCALAGLAAALPLASVGEYGAAPLCLPVAPREGOPALGFTVALVMNSFCEPLV 747
 DB 540 AIMGWGWCCFLMLLPVIGISSYAKVSIQPM--DFEPLALAVIILVLLNIVAFIIV 597
 QY 748 AGAVITKLYCDLPRDDE-AWWDCAVNRHVAMLIRADGLLYCPVAFSLASMLGEPVTP 806
 DB 598 CSCYVKTITTYRNPQYNDGDKDTIARMAVLIPTDFCMAPISFYALSALMNKPLITVT 657
 QY 807 AVKSVLLVLPPLACLPNPLLYLFNPHFRDRLRLRPAGDSGLAYAAAG----- 857
 DB 658 NSKILLVLEVPNLANCPPLIATFTKAFQRDVFIILSKFGICKRQAOAYRQGRVSPKNSA 717
 QY 858 --ELEKSSCSTQALVAFSDVDLILEAS 883
 DB 718 GLOIQKVTBDMROSLPMMODEYELLENS 745
 RESULT 11
 TSHR_MOUSE
 ID TSHR_MOUSE STANDARD; PRT; 764 AA.
 AC P47750;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone
 DE receptor).
 GN TSHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE=Thyroid;
 RX MEDLINE=9422423; PubMed=8170469;
 RA Stein S.A., Oates E.L., Hall C.R., Grumbles R.M., Fernandez L.M.,
 Taylor N.A., Puett D., Jin S.;
 RT "Identification of a point mutation in the thyrotropin receptor of
 the hyl/hyl hypothyroid mouse";
 RL Mol. Endocrinol. 8:129-138(1994).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
 CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
 CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DISEASE: DEFECTS IN TSHR ARE THE CAUSE OF HYP/HYP HYPOTHYROIDISM.
 CC AN AUTOSOMAL RECESSIVE, FETAL-ONSET, SEVERE HYPOTHYROIDISM RELATED
 TO TSH HYPORESponsiveness AND ASSOCIATED WITH ELEVATED TSH.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 8 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U02601; AAA53209.1; -;
 DR EMBL: U02602; AAB0455.1; -;
 DR HSSP: P16473; 1XUM.
 DR GCRdb: GCR_1718; -;
 DR GCRdb: GCR_1719; -;
 DR MGD: MGI:98849; TSHr.
 DR InterPro: IPR00276; GPCR_Rhodpsn.
 DR InterPro: IPR01611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR PRINTS: PR00373; GLYCHOMONER.
 DR PRINTS: PR01145; TSHRECEPTOR.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Disease mutation.
 FT SIGNAL 1 21
 FT CHAIN 22 764
 FT DOMAIN 22 413
 FT TRANSMEM 414 441
 FT DOMAIN 442 450
 FT TRANSMEM 451 473
 FT DOMAIN 474 494
 FT TRANSMEM 495 517
 FT DOMAIN 518 537
 FT TRANSMEM 538 560
 FT DOMAIN 561 580
 FT TRANSMEM 581 602
 FT DOMAIN 603 625
 FT TRANSMEM 626 649
 FT DOMAIN 650 660
 FT TRANSMEM 661 682
 FT DOMAIN 683 764
 FT REPEAT 51 74
 FT REPEAT 100 124
 FT REPEAT 125 150
 FT REPEAT 151 174
 FT REPEAT 176 199
 FT REPEAT 200 223
 FT REPEAT 225 248
 FT REPEAT 264 288
 FT DISULFID 494 569
 FT CARBOHYD 77 77
 FT CARBOHYD 99 99
 FT CARBOHYD 177 177
 FT CARBOHYD 198 198

FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 556 556 P -> L (IN HYPOTHYROIDISM).
 SQ SEQUENCE 764 AA; 86583 MW; 6ADD2CC72P018317 CRC64;
 Query Match 9.6%; Score 486; DB 1; Length 764;
 Best Local Similarity 21.7%; Pred. No. 7.6e-26;
 Matches 211; Conservative 132; Mismatches 346; Indels 284; Gaps 30;
 QY 4 PGLRLMLCALACARRAGAPQPPGPTACAPCHC-QEEDGIMLSADCELGISAVPG 62
 DB 3 PGLLLLVLLALSRIR-----GKECAPPECHQEDDFRT-CREL----- 44
 QY 63 DLDPLTAYIDLNNMLTEPLGFLHRLPLELRISGNHLSHPGQAFSGLSYSLTIMQ 122
 DB 45 -----HRIPLSPS-----TQTLKLETHKLTPLSLAFSLPNNISYIS 84
 QY 123 NNOIGCPALMELPSLOLRDANMLISVPRFEGSLSLH-----WLDNA 173
 DB 85 -----IDATLQRLP-HSFYNSKMHIELRMTSLTFTIDPPA 121
 QY 174 LTEIP-----VRALNLPALQAMTLAL-----NRISHIDYAFQNL-TSLVY 214
 DB 122 LTELPLKFLGIFNTGRIPLDKIYSTDIFLEITDNPVTSVENAFQGLCHETILT 181
 QY 215 LHLHNNRIQHLGTHSEGLHNLLETLDNYKQLQEPFAITLRLQELGHHNNKAIPE 274
 DB 182 LKLYNNGFTSYOGHAFNGT-KLDAYVLLNKNY-----LRAIDN 218
 QY 275 KAFGNLLOTTHPYDNP-IOEVGRSAFOYLPR-----LHLSINGAMDIOEPDLKGT 328
 DB 219 DAFGG-----YSGPTLIDVSTSVTALPSGHEHLKELAKDWTLKLPL-----L 265
 QY 329 SLEILITLRAGIRLLPSGMCQQLRLRYVLELSHQIIEPLSLHRCQLEETIGLOHRIWE 388
 DB 266 SLSPFLHLTRADLS-YPSHCACFKNQKKIRGLIESIMKNESIRNLORKSVNLRGPYQ 324
 QY 389 -----IGATPFGSLSSICALDLSWMNIRSHPEAFSTLSLYKLDLTDLNQTLLP 438
 DB 325 EYEDPDGDSYG---YKNSKFKQ-----ESPSNSHYV----- 354
 QY 439 IAGLGIMHLKGLMALSQAFSKDSPEPKIRILEVPYAYOCPPGMCASFPGKSGOMEAE 498
 DB 355 -----FE-----EOE 360
 QY 499 DLHIDDESSKRPGLARQAEHNHYDDLDQLMEDSKRHPVQCSPPRGPKPEYL 558
 DB 361 DEVVFGQELKNPQRETLQAFESHYDYV-----CGNEDMVCPTPSDEFNPECDI 411
 QY 559 PESMGIRLAVMAIVLSVINGVLYLVFAGGAPALPVPKVVGAIGANTLTISGGL 618
 DB 412 MGYRFLRIYVWFVSLALAGNIFVLLIILTSYKLTVE-RLMCMNLAFADFCMGVYLLI 470
 QY 619 ASVDALITGQFSEYGAKEWTGLGCRATGFLAVLGSEASVLLITLAOVCSVSCVARYG 678
 DB 471 ASVDLYTHSEYVNHAIIDMOTGPGCNTAGFFVFASELVLYLTTLTTRWATTFAMRLD 530
 QY 679 KSPSLGSVAVGLGCLALAGLAALPLASVGEYASPCLPAPPEGPPALAGFTVALM 738
 DB 531 KRIIRLHAYTTIAGWVSCFLLALPLMWGISYAVSICLPR--DTDPLALAVLVVL 588
 QY 739 MNSFCFLVAGAVYIKLYCDL-----PRGDFEAVNDCAVRHVAVMLFADGLLYPVAVL 792
 DB 589 LNVAVFVVCSCYKITYIVRNPNQYNRDK-----DTKIAKRAVALITFDPCMAPISEY 643
 QY 793 SFASMTGLFPVTPPEAVKSVLLVYLPPLCNPLNPLVILFNPHFRDRLRRPAPAGSGPLA 852
 DB 644 ALSALNNKPLIYVNSKILLVLEPLNSCANPFLAIFKAFQROVFTILSKFGICKRQA 703
 QY 853 YAAAG-----ELKSSCDSTQALVAFSDVLLILEASEAGRPGLFTYGPFSVTL 901
 DB 704 QATQGRVCPNNSTGIQIKIPDTRQSLPNNQDYTELLGNSQL----- 747
 QY 902 ISCOGPGAPRLLEG 914

```

DB 748 -----APKLOG 753
||:|:|
RESULT 12
TSR-BOVIN STANDARD; PRT; 763 AA.
AC Q27987;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thyrotropin receptor precursor (TSR-R) (Thyroid stimulating hormone
DE receptor).
CN TSR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RA MEDLINE=97280128; PubMed=9134497;
RA Silverides D.W., Houde A., Ethier J.F., Lussier J.G.;
RT "Bovine thyrotropin receptor cDNA is characterized by full-length and
RT truncated transcripts."
RL J. Mol. Endocrinol. 18:101-112(1997).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
CC CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
CC CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSR SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U15570; AAC18639.1; -
CC HSSP: P16473; 1XDM.
CC
CC GCRDB: GCR_1225; -
CC InterPro: IPR000276; GPCR_Rhodopsin.
CC InterPro: IPR001611; LRR.
CC Pfam: PF00001; 7tm.1; 1.
CC Pfam: PF00560; LRR; 3.
CC PRINTS: PR00373; GLYCHROMONER.
CC PRINTS: PR01145; TSRRECEPTOR.
CC PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
CC PROSITE: PS50162; G_PROTEIN_RECPT_FL_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 21
FT CHAIN 22 763
FT DOMAIN 22 412
FT TRANSMEM 413 440
FT DOMAIN 441 449
FT TRANSMEM 450 472
FT DOMAIN 473 493
FT TRANSMEM 494 516
FT DOMAIN 517 536
FT TRANSMEM 537 559
FT DOMAIN 560 579
FT TRANSMEM 580 601
FT DOMAIN 602 624
FT TRANSMEM 625 648
FT DOMAIN 649 659
FT TRANSMEM 660 681

```

```

FT DOMAIN 662 763
FT REPEAT 51 74
FT REPEAT 125 150
FT REPEAT 151 174
FT REPEAT 176 199
FT REPEAT 201 223
FT REPEAT 225 248
FT DISULFID 493 568
FT CARBOHYD 77 77
FT CARBOHYD 99 99
FT CARBOHYD 177 177
FT CARBOHYD 198 198
FT CARBOHYD 302 302
SQ SEQUENCE 763 AA; 86431 MM; 3589647 EDTA8C CRC64;
Query Match 9.5%; Score 479; DB 1; Length 763;
Best Local Similarity 23.1%; Pred. No. 2,3e-25;
Matches 210; Conservativity 139; Mismatches 321; Indels 238; Gaps 36;
QY 31 GPTACPA-PCHC-QEDGIMLSADCELSGSAVPGDLPFLAYLDSMNNLTRELQGLFHH 88
DB 20 GGERCPSPCECROEDDEPRVT--CKD--IQSIP-SLEPST-----QT 56
QY 89 LRFELRLSGNHLSHIPGAFSGLYSILKILMLQNNOLGIGTAEALWELPSLQSLRDAN 148
DB 57 LKFI-----THLKTISRAPSNLPNISRYLS-----IDAT 88
QY 149 LISLVERSEFGLSLRLHML-DONALTEIPYALNMLPALQAMTLALNRIHIPVAFQ 207
DB 89 LQOL-ESHSEFYNLSKVTHIEIRNTRSLTYIDSGALKEPLKFLGIFNTGLRVFPD----- 143
QY 208 NLTSL-----VVLHLHNN-RIGHGTSFEGINLHLETLNLYKNLQEPVAIRTLGRLO 260
DB 144 -LTKIYSDVFFLEITDNPYMTSIPANAFQGLCN-ETLTLC----- 183
QY 261 ELGHNNNKAIPEKAFGNPDLQTHFYDNP-IOFVRSAPFOYLPKHLTSLINGANDIQ 319
DB 184 ---LYNNFGTSIQGHAFNGTK-LDAVYINKKYLTVIGQDAF----- 221
QY 320 EFPPLKGTSLIELTLTRAGIRLPLSGMCOQLPRLVLEISHNOIEELPSLRQKLEEI 379
DB 222 -----AGVYSGPT-----LSDISYTSVTALPSK----- 244
QY 380 GLQHNRIWEIGADTFSSQLSLQALDLSMNAIRSIHPEAFSLHLYNKDLF-----D 431
DB 245 GLEH-----LKEILIAN-TW-TLRKL-PLSLFLH-LTRADLSYSHCCAFK 287
QY 432 NQTLTDLPLAGIGIMHLKLGKGNLALSOAFKSDSPKLR-----ILEVYAYQCCPYGMC 485
DB 288 NQ-----KIRGILO-SIMCNESSIRGLRKRKSASALNGPFYGEYEDIGDG 332
QY 486 ASPEKASGOW-----EADLHLHDESSKRPGLGLARAEHNYDODLELO 531
DB 333 SAGYKENSKEQDQOSNHYVFEEOQEDETIGFQOQLRNQEETLOAFDHYDTVAGGS 392
QY 532 LEMEDSKPHSVOCSPPTGPKPCDEYLFESWGIRLAWAIVLTVLCNGVLLTVFAGGP 591
DB 393 EDM-----VCTPKSDFNFCEDIMGYKFLRIYVWFSLIALLGNGVYVYLLTSY 443
QY 592 APPLPYKVVYGAALAGANTLTGSCGLASVDALTFGGFSEYGARWETGLCGRAFLAVL 651
DB 444 KLVLP-REFLNCNLAFADFCMGLVLLIASVDLTOSYRYNHAIDMOTGPCNAGFEYV 502
QY 652 GSEASVYLITLAAVQCSVSVCYRAYGKSPSLGVSAGVAGLALAGLAAVLAAGEV 711
DB 503 ASLSTYTLVTLVTLLEHMHATTFAMRDKIRLHAIVYIMGVCVCFLLALPLVGISSV 562
QY 712 GASPLCLPYAPPEGOPALGFTVALVMMNSFCELVVAGAYIKLYCDLP-----GDPEAV 766
DB 563 AKVSIOLPVM-DTEPLALAYIILVLLNIAFIIVCACYVKIYITVRNPHVPGD----- 616
QY 767 WDCAMVRYHVAWLIIFDGLLYCPAFLSFASMGCLFVYTEAKSVLLVVLPLPACINPLL 826

```

Db 617 KDRIRARMAVLFTDWCAPISFYALSALMKNPLTVNTSKILLVFLPNSCANPFL 676
 QY 827 YLLEPNHFRDRLRRPRAGDSGLAYAAAGE-----LEKSSCSTOALVAFSD 875
 Db 677 YALFTKAFQRDVFEFLSKFGICKRQAOAYRGORVSPKNSGTGIRQVKVPPDVROSLPNVOD 736
 QY 876 VDLILENS 883
 Db 737 DYELLENS 744

RESULT 13
 LSHR_RAT STANDARD; PRT: 700 AA.
 AC P16235; P70646; Q63807; Q63808; Q63809;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
 GN (LH-R) (Luteinizing hormone receptor).
 OS LHCGR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RX SEQUENCE FROM N.A.,
 RA MEDLINE=89332512; PubMed=2502842;
 RA McFarland K.C., Sprengel R., Phillips H.S., Koehler M.,
 RA Rosemblit N., Nikolic K., Segaloff D.L., Seeburg P.H.;
 RT "Lutropin-choriogonadotropic receptor: an unusual member of the G
 RL protein-coupled receptor family.";
 RL Science 245:494-499(1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STAIN-SPRAGUE-DAWLEY; TISSUE-Ovary;
 RX MEDLINE=92347604; PubMed=1353463;
 RA Aatsinki J.T., Pietila E.M., Lakkakorpi J.T., Rajanemi H.J.;
 RT "Expression of the LH/CG receptor gene in rat ovarian tissue is
 RT regulated by an extensive alternative splicing of the primary
 RT transcript.";
 RL Mol. Cell. Endocrinol. 84:127-135(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91209270; PubMed=2019252;
 RA Koo Y.B., Slaughter R.G., Ji T.H.;
 RT "Structure of the luteinizing hormone receptor gene and multiple
 RL exons of the coding sequence.";
 RL Endocrinology 128:2297-2308(1991).
 RN [4]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=91006819; PubMed=1976554;
 RA Bernard M.P., Myers R.V., Moyle W.R.;
 RT "Cloning of rat lutropin (LH) receptor analogs lacking the soybean
 RL lectin domain.";
 RL Mol. Cell. Endocrinol. 71:R19-R23(1990).
 RN [5]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=91126285; PubMed=2281186;
 RA Segaloff D.L., Sprengel R., Nikolic K., Ascoli M.;
 RT "Structure of the lutropin/choriogonadotropin receptor.";
 RL Recent Prog. Horm. Res. 46:261-303(1990).
 RN [6]
 RP SEQUENCE OF 295-700 FROM N.A.
 RX MEDLINE=91060531; PubMed=2174034;
 RA Tsai-Morris C.H., Buczek E., Wang W., Dufan M.L.;
 RT "Intronic nature of the rat luteinizing hormone receptor gene defines
 RL a soluble receptor subspecies with hormone binding activity.";
 RL J. Biol. Chem. 265:19385-19388(1990).
 RN [7]
 RP SEQUENCE OF 27-37.
 RX MEDLINE=89174723; PubMed=2925659;
 RA Roche P.C., Ryan R.J.;

RT "Purification, characterization, and amino-terminal sequence of rat
 RT ovarian receptor for luteinizing hormone/human choriongonadotropin.";
 RL J. Biol. Chem. 264:4638-4641(1989).
 RN [8]
 RP MEDLINE=91332007; PubMed=1714448;
 RX MOTAGENESIS.
 RA Ji I., Ji T.H.;
 RT "Asp383 in the second transmembrane domain of the lutropin receptor
 RT is important for high affinity hormone binding and cAMP production.";
 RL J. Biol. Chem. 266:14953-14957(1991).
 CC -1- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
 CC ACTIVATE ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 11 ISOFORMS WHICH DIFFER IN
 CC SUBCELLULAR LOCATION ARE PRODUCED BY ALTERNATIVE SPLICING
 CC OF THE SAME GENE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----

DR EMBL; M26199; AAA41528.1; -;
 DR EMBL; M61212; AAA41527.1; -;
 DR EMBL; M61211; AAA41527.1; JOINED.
 DR EMBL; S40803; AAB22680.1; -;
 DR EMBL; S40787; AAB22680.1; JOINED.
 DR EMBL; S40903; AAB22680.1; JOINED.
 DR EMBL; S40904; AAB22680.1; JOINED.
 DR EMBL; S40905; AAB22680.1; JOINED.
 DR EMBL; S40906; AAB22680.1; JOINED.
 DR EMBL; S40907; AAB22680.1; JOINED.
 DR EMBL; S40908; AAB22680.1; JOINED.
 DR EMBL; S40909; AAB22680.1; JOINED.
 DR EMBL; S40910; AAB22680.1; JOINED.
 DR EMBL; S40911; AAB22680.1; JOINED.
 DR EMBL; S40912; AAB22680.1; JOINED.
 DR EMBL; S40913; AAB22680.1; JOINED.
 DR EMBL; S40914; AAB22680.1; JOINED.
 DR EMBL; S40915; AAB22680.1; JOINED.
 DR EMBL; S40916; AAB22680.1; JOINED.
 DR EMBL; S40917; AAB22680.1; JOINED.
 DR EMBL; S40918; AAB22680.1; JOINED.
 DR EMBL; S40919; AAB22680.1; JOINED.
 DR EMBL; S40920; AAB22680.1; JOINED.
 DR EMBL; S40921; AAB22680.1; JOINED.
 DR EMBL; S40922; AAB22680.1; JOINED.
 DR EMBL; S40923; AAB22680.1; JOINED.
 DR EMBL; S40924; AAB22680.1; JOINED.
 DR EMBL; S40925; AAB22680.1; JOINED.
 DR EMBL; S40926; AAB22680.1; JOINED.
 DR EMBL; S40927; AAB22680.1; JOINED.
 DR EMBL; S40928; AAB22680.1; JOINED.
 DR EMBL; S40929; AAB22680.1; JOINED.
 DR EMBL; S40930; AAB22680.1; JOINED.
 DR EMBL; S40931; AAB22680.1; JOINED.
 DR EMBL; S40932; AAB22680.1; JOINED.
 DR EMBL; S40933; AAB22680.1; JOINED.
 DR EMBL; S40934; AAB22680.1; JOINED.
 DR EMBL; S40935; AAB22680.1; JOINED.
 DR EMBL; S40936; AAB22680.1; JOINED.
 DR EMBL; S40937; AAB22680.1; JOINED.
 DR EMBL; S40938; AAB22680.1; JOINED.
 DR EMBL; S40939; AAB22680.1; JOINED.
 DR EMBL; S40940; AAB22680.1; JOINED.
 DR EMBL; S40941; AAB22680.1; JOINED.
 DR EMBL; S40942; AAB22680.1; JOINED.
 DR EMBL; S40943; AAB22680.1; JOINED.
 DR EMBL; S40944; AAB22680.1; JOINED.
 DR EMBL; S40945; AAB22680.1; JOINED.
 DR EMBL; S40946; AAB22680.1; JOINED.
 DR EMBL; S40947; AAB22680.1; JOINED.
 DR EMBL; S40948; AAB22680.1; JOINED.
 DR EMBL; S40949; AAB22680.1; JOINED.
 DR EMBL; S40950; AAB22680.1; JOINED.
 DR EMBL; S40951; AAB22680.1; JOINED.
 DR EMBL; S40952; AAB22680.1; JOINED.
 DR EMBL; S40953; AAB22680.1; JOINED.
 DR EMBL; S40954; AAB22680.1; JOINED.
 DR EMBL; S40955; AAB22680.1; JOINED.
 DR EMBL; S40956; AAB22680.1; JOINED.
 DR EMBL; S40957; AAB22680.1; JOINED.
 DR EMBL; S40958; AAB22680.1; JOINED.
 DR EMBL; S40959; AAB22680.1; JOINED.
 DR EMBL; S40960; AAB22680.1; JOINED.
 DR EMBL; S40961; AAB22680.1; JOINED.
 DR EMBL; S40962; AAB22680.1; JOINED.
 DR EMBL; S40963; AAB22680.1; JOINED.
 DR EMBL; S40964; AAB22680.1; JOINED.
 DR EMBL; S40965; AAB22680.1; JOINED.
 DR EMBL; S40966; AAB22680.1; JOINED.
 DR EMBL; S40967; AAB22680.1; JOINED.
 DR EMBL; S40968; AAB22680.1; JOINED.
 DR EMBL; S40969; AAB22680.1; JOINED.
 DR EMBL; S40970; AAB22680.1; JOINED.
 DR EMBL; S40971; AAB22680.1; JOINED.
 DR EMBL; S40972; AAB22680.1; JOINED.
 DR EMBL; S40973; AAB22680.1; JOINED.
 DR EMBL; S40974; AAB22680.1; JOINED.
 DR EMBL; S40975; AAB22680.1; JOINED.
 DR EMBL; S40976; AAB22680.1; JOINED.
 DR EMBL; S40977; AAB22680.1; JOINED.
 DR EMBL; S40978; AAB22680.1; JOINED.
 DR EMBL; S40979; AAB22680.1; JOINED.
 DR EMBL; S40980; AAB22680.1; JOINED.
 DR EMBL; S40981; AAB22680.1; JOINED.
 DR EMBL; S40982; AAB22680.1; JOINED.
 DR EMBL; S40983; AAB22680.1; JOINED.
 DR EMBL; S40984; AAB22680.1; JOINED.
 DR EMBL; S40985; AAB22680.1; JOINED.
 DR EMBL; S40986; AAB22680.1; JOINED.
 DR EMBL; S40987; AAB22680.1; JOINED.
 DR EMBL; S40988; AAB22680.1; JOINED.
 DR EMBL; S40989; AAB22680.1; JOINED.
 DR EMBL; S40990; AAB22680.1; JOINED.
 DR EMBL; S40991; AAB22680.1; JOINED.
 DR EMBL; S40992; AAB22680.1; JOINED.
 DR EMBL; S40993; AAB22680.1; JOINED.
 DR EMBL; S40994; AAB22680.1; JOINED.
 DR EMBL; S40995; AAB22680.1; JOINED.
 DR EMBL; S40996; AAB22680.1; JOINED.
 DR EMBL; S40997; AAB22680.1; JOINED.
 DR EMBL; S40998; AAB22680.1; JOINED.
 DR EMBL; S40999; AAB22680.1; JOINED.
 DR EMBL; S41000; AAB22680.1; JOINED.
 DR EMBL; S41001; AAB22680.1; JOINED.
 DR EMBL; S41002; AAB22680.1; JOINED.
 DR EMBL; S41003; AAB22680.1; JOINED.
 DR EMBL; S41004; AAB22680.1; JOINED.
 DR EMBL; S41005; AAB22680.1; JOINED.
 DR EMBL; S41006; AAB22680.1; JOINED.
 DR EMBL; S41007; AAB22680.1; JOINED.
 DR EMBL; S41008; AAB22680.1; JOINED.
 DR EMBL; S41009; AAB22680.1; JOINED.
 DR EMBL; S41010; AAB22680.1; JOINED.
 DR EMBL; S41011; AAB22680.1; JOINED.
 DR EMBL; S41012; AAB22680.1; JOINED.
 DR EMBL; S41013; AAB22680.1; JOINED.
 DR EMBL; S41014; AAB22680.1; JOINED.
 DR EMBL; S41015; AAB22680.1; JOINED.
 DR EMBL; S41016; AAB22680.1; JOINED.
 DR EMBL; S41017; AAB22680.1; JOINED.
 DR EMBL; S41018; AAB22680.1; JOINED.
 DR EMBL; S41019; AAB22680.1; JOINED.
 DR EMBL; S41020; AAB22680.1; JOINED.
 DR EMBL; S41021; AAB22680.1; JOINED.
 DR EMBL; S41022; AAB22680.1; JOINED.
 DR EMBL; S41023; AAB22680.1; JOINED.
 DR EMBL; S41024; AAB22680.1; JOINED.
 DR EMBL; S41025; AAB22680.1; JOINED.
 DR EMBL; S41026; AAB22680.1; JOINED.
 DR EMBL; S41027; AAB22680.1; JOINED.
 DR EMBL; S41028; AAB22680.1; JOINED.
 DR EMBL; S41029; AAB22680.1; JOINED.
 DR EMBL; S41030; AAB22680.1; JOINED.
 DR EMBL; S41031; AAB22680.1; JOINED.
 DR EMBL; S41032; AAB22680.1; JOINED.
 DR EMBL; S41033; AAB22680.1; JOINED.
 DR EMBL; S41034; AAB22680.1; JOINED.
 DR EMBL; S41035; AAB22680.1; JOINED.
 DR EMBL; S41036; AAB22680.1; JOINED.
 DR EMBL; S41037; AAB22680.1; JOINED.
 DR EMBL; S41038; AAB22680.1; JOINED.
 DR EMBL; S41039; AAB22680.1; JOINED.
 DR EMBL; S41040; AAB22680.1; JOINED.
 DR EMBL; S41041; AAB22680.1; JOINED.
 DR EMBL; S41042; AAB22680.1; JOINED.
 DR EMBL; S41043; AAB22680.1; JOINED.
 DR EMBL; S41044; AAB22680.1; JOINED.
 DR EMBL; S41045; AAB22680.1; JOINED.
 DR EMBL; S41046; AAB22680.1; JOINED.
 DR EMBL; S41047; AAB22680.1; JOINED.
 DR EMBL; S41048; AAB22680.1; JOINED.
 DR EMBL; S41049; AAB22680.1; JOINED.
 DR EMBL; S41050; AAB22680.1; JOINED.
 DR EMBL; S41051; AAB22680.1; JOINED.
 DR EMBL; S41052; AAB22680.1; JOINED.
 DR EMBL; S41053; AAB22680.1; JOINED.
 DR EMBL; S41054; AAB22680.1; JOINED.
 DR EMBL; S41055; AAB22680.1; JOINED.
 DR EMBL; S41056; AAB22680.1; JOINED.
 DR EMBL; S41057; AAB22680.1; JOINED.
 DR EMBL; S41058; AAB22680.1; JOINED.
 DR EMBL; S41059; AAB22680.1; JOINED.
 DR EMBL; S41060; AAB22680.1; JOINED.
 DR EMBL; S41061; AAB22680.1; JOINED.
 DR EMBL; S41062; AAB22680.1; JOINED.
 DR EMBL; S41063; AAB22680.1; JOINED.
 DR EMBL; S41064; AAB22680.1; JOINED.
 DR EMBL; S41065; AAB22680.1; JOINED.
 DR EMBL; S41066; AAB22680.1; JOINED.
 DR EMBL; S41067; AAB22680.1; JOINED.
 DR EMBL; S41068; AAB22680.1; JOINED.
 DR EMBL; S41069; AAB22680.1; JOINED.
 DR EMBL; S41070; AAB22680.1; JOINED.
 DR EMBL; S41071; AAB22680.1; JOINED.
 DR EMBL; S41072; AAB22680.1; JOINED.
 DR EMBL; S41073; AAB22680.1; JOINED.
 DR EMBL; S41074; AAB22680.1; JOINED.
 DR EMBL; S41075; AAB22680.1; JOINED.
 DR EMBL; S41076; AAB22680.1; JOINED.
 DR EMBL; S41077; AAB22680.1; JOINED.
 DR EMBL; S41078; AAB22680.1; JOINED.
 DR EMBL; S41079; AAB22680.1; JOINED.
 DR EMBL; S41080; AAB22680.1; JOINED.
 DR EMBL; S41081; AAB22680.1; JOINED.
 DR EMBL; S41082; AAB22680.1; JOINED.
 DR EMBL; S41083; AAB22680.1; JOINED.
 DR EMBL; S41084; AAB22680.1; JOINED.
 DR EMBL; S41085; AAB22680.1; JOINED.
 DR EMBL; S41086; AAB22680.1; JOINED.
 DR EMBL; S41087; AAB22680.1; JOINED.
 DR EMBL; S41088; AAB22680.1; JOINED.
 DR EMBL; S41089; AAB22680.1; JOINED.
 DR EMBL; S41090; AAB22680.1; JOINED.
 DR EMBL; S41091; AAB22680.1; JOINED.
 DR EMBL; S41092; AAB22680.1; JOINED.
 DR EMBL; S41093; AAB22680.1; JOINED.
 DR EMBL; S41094; AAB22680.1; JOINED.
 DR EMBL; S41095; AAB22680.1; JOINED.
 DR EMBL; S41096; AAB22680.1; JOINED.
 DR EMBL; S41097; AAB22680.1; JOINED.
 DR EMBL; S41098; AAB22680.1; JOINED.
 DR EMBL; S41099; AAB22680.1; JOINED.
 DR EMBL; S41100; AAB22680.1; JOINED.
 DR EMBL; S41101; AAB22680.1; JOINED.
 DR EMBL; S41102; AAB22680.1; JOINED.
 DR EMBL; S41103; AAB22680.1; JOINED.
 DR EMBL; S41104; AAB22680.1; JOINED.
 DR EMBL; S41105; AAB22680.1; JOINED.
 DR EMBL; S41106; AAB22680.1; JOINED.
 DR EMBL; S41107; AAB22680.1; JOINED.
 DR EMBL; S41108; AAB22680.1; JOINED.
 DR EMBL; S41109; AAB22680.1; JOINED.
 DR EMBL; S41110; AAB22680.1; JOINED.
 DR EMBL; S41111; AAB22680.1; JOINED.
 DR EMBL; S41112; AAB22680.1; JOINED.
 DR EMBL; S41113; AAB22680.1; JOINED.
 DR EMBL; S41114; AAB22680.1; JOINED.
 DR EMBL; S41115; AAB22680.1; JOINED.
 DR EMBL; S41116; AAB22680.1; JOINED.
 DR EMBL; S41117; AAB22680.1; JOINED.
 DR EMBL; S41118; AAB22680.1; JOINED.
 DR EMBL; S41119; AAB22680.1; JOINED.
 DR EMBL; S41120; AAB22680.1; JOINED.
 DR EMBL; S41121; AAB22680.1; JOINED.
 DR EMBL; S41122; AAB22680.1; JOINED.
 DR EMBL; S41123; AAB22680.1; JOINED.
 DR EMBL; S41124; AAB22680.1; JOINED.
 DR EMBL; S41125; AAB22680.1; JOINED.
 DR EMBL; S41126; AAB22680.1; JOINED.
 DR EMBL; S41127; AAB22680.1; JOINED.
 DR EMBL; S41128; AAB22680.1; JOINED.
 DR EMBL; S41129; AAB22680.1; JOINED.
 DR EMBL; S41130; AAB22680.1; JOINED.
 DR EMBL; S41131; AAB22680.1; JOINED.
 DR EMBL; S41132; AAB22680.1; JOINED.
 DR EMBL; S41133; AAB22680.1; JOINED.
 DR EMBL; S41134; AAB22680.1; JOINED.
 DR EMBL; S41135; AAB22680.1; JOINED.
 DR EMBL; S41136; AAB22680.1; JOINED.
 DR EMBL; S41137; AAB22680.1; JOINED.
 DR EMBL; S41138; AAB22680.1; JOINED.
 DR EMBL; S41139; AAB22680.1; JOINED.
 DR EMBL; S41140; AAB22680.1; JOINED.
 DR EMBL; S41141; AAB22680.1; JOINED.
 DR EMBL; S41142; AAB22680.1; JOINED.
 DR EMBL; S41143; AAB22680.1; JOINED.
 DR EMBL; S41144; AAB22680.1; JOINED.
 DR EMBL; S41145; AAB22680.1; JOINED.
 DR EMBL; S41146; AAB22680.1; JOINED.
 DR EMBL; S41147; AAB22680.1; JOINED.
 DR EMBL; S41148; AAB22680.1; JOINED.
 DR EMBL; S41149; AAB22680.1; JOINED.
 DR EMBL; S41150; AAB22680.1; JOINED.
 DR EMBL; S41151; AAB22680.1; JOINED.
 DR EMBL; S41152; AAB22680.1; JOINED.
 DR EMBL; S41153; AAB22680.1; JOINED.
 DR EMBL; S41154; AAB22680.1; JOINED.
 DR EMBL; S41155; AAB22680.1; JOINED.
 DR EMBL; S41156; AAB22680.1; JOINED.
 DR EMBL; S41157; AAB22680.1; JOINED.
 DR EMBL; S41158; AAB22680.1; JOINED.
 DR EMBL; S41159; AAB22680.1; JOINED.
 DR EMBL; S41160; AAB22680.1; JOINED.
 DR EMBL; S41161; AAB22680.1; JOINED.
 DR EMBL; S41162; AAB22680.1; JOINED.
 DR EMBL; S41163; AAB22680.1; JOINED.
 DR EMBL; S41164; AAB22680.1; JOINED.
 DR EMBL; S41165; AAB22680.1; JOINED.
 DR EMBL; S41166; AAB22680.1; JOINED.
 DR EMBL; S41167; AAB22680.1; JOINED.
 DR EMBL; S41168; AAB22680.1; JOINED.
 DR EMBL; S41169; AAB22680.1; JOINED.
 DR EMBL; S41170; AAB22680.1; JOINED.
 DR EMBL; S41171; AAB22680.1; JOINED.
 DR EMBL; S41172; AAB22680.1; JOINED.
 DR EMBL; S41173; AAB22680.1; JOINED.
 DR EMBL; S41174; AAB22680.1; JOINED.
 DR EMBL; S41175; AAB22680.1; JOINED.
 DR EMBL; S41176; AAB22680.1; JOINED.
 DR EMBL; S41177; AAB22680.1; JOINED.
 DR EMBL; S41178; AAB22680.1; JOINED.
 DR EMBL; S41179; AAB22680.1; JOINED.
 DR EMBL; S41180; AAB22680.1; JOINED.
 DR EMBL; S41181; AAB22680.1; JOINED.
 DR EMBL; S41182; AAB22680.1; JOINED.
 DR EMBL; S41183; AAB22680.1; JOINED.
 DR EMBL; S41184; AAB22680.1; JOINED.
 DR EMBL; S41185; AAB22680.1; JOINED.
 DR EMBL; S41186; AAB22680.1; JOINED.
 DR EMBL; S41187; AAB22680.1; JOINED.
 DR EMBL; S41188; AAB22680.1; JOINED.
 DR EMBL; S41189; AAB22680.1; JOINED.
 DR EMBL; S41190; AAB22680.1; JOINED.
 DR EMBL; S41191; AAB22680.1; JOINED.
 DR EMBL; S41192; AAB22680.1; JOINED.
 DR EMBL; S41193; AAB22680.1; JOINED.
 DR EMBL; S41194; AAB22680.1; JOINED.
 DR EMBL; S41195; AAB22680.1; JOINED.
 DR EMBL; S41196; AAB22680.1; JOINED.
 DR EMBL; S41197; AAB22680.1; JOINED.
 DR EMBL; S41198; AAB22680.1; JOINED.
 DR EMBL; S41199; AAB22680.1; JOINED.
 DR EMBL; S41200; AAB22680.1; JOINED.
 DR EMBL; S41201; AAB22680.1; JOINED.
 DR EMBL; S41202; AAB22680.1; JOINED.
 DR EMBL; S41203; AAB22680.1; JOINED.
 DR EMBL; S41204; AAB22680.1; JOINED.
 DR EMBL; S41205; AAB22680.1; JOINED.
 DR EMBL; S41206; AAB22680.1; JOINED.
 DR EMBL; S41207; AAB22680.1; JOINED.
 DR EMBL; S41208; AAB22680.1; JOINED.
 DR EMBL; S41209; AAB22680.1; JOINED.
 DR EMBL; S41210; AAB22680.1; JOINED.
 DR EMBL; S41211; AAB22680.1; JOINED.
 DR EMBL; S41212; AAB22680.1; JOINED.
 DR EMBL; S41213; AAB22680.1; JOINED.
 DR EMBL; S41214; AAB22680.1; JOINED.
 DR EMBL; S41215; AAB22680.1; JOINED.
 DR EMBL; S41216; AAB22680.1; JOINED.
 DR EMBL; S41217; AAB22680.1; JOINED.
 DR EMBL; S41218; AAB22680.1; JOINED.
 DR EMBL; S41219; AAB22680.1; JOINED.
 DR EMBL; S41220; AAB22680.1; JOINED.
 DR EMBL; S41221; AAB22680.1; JOINED.
 DR EMBL; S41222; AAB22680.1; JOINED.
 DR EMBL; S41223; AAB22680.1; JOINED.
 DR EMBL; S41224; AAB22680.1; JOINED.
 DR EMBL; S41225; AAB22680.1; JOINED.
 DR EMBL; S41226; AAB22680.1; JOINED.
 DR EMBL; S41227; AAB22680.1; JOINED.
 DR EMBL; S41228; AAB22680.1; JOINED.
 DR EMBL; S41229; AAB22680.1; JOINED.
 DR EMBL; S41230; AAB22680.1; JOINED.
 DR EMBL; S41231; AAB22680.1; JOINED.
 DR EMBL; S41232; AAB22680.1; JOINED.
 DR EMBL; S41233; AAB22680.1; JOINED.
 DR EMBL; S41234; AAB22680.1; JOINED.
 DR EMBL; S41235; AAB22680.1; JOINED.
 DR EMBL; S41236; AAB22680.1; JOINED.
 DR EMBL; S41237; AAB22680.1; JOINED.
 DR EMBL; S41238; AAB22680.1; JOINED.
 DR EMBL; S41239; AAB22680.1; JOINED.
 DR EMBL; S41240; AAB22680.1; JOINED.
 DR EMBL; S41241; AAB22680.1; JOINED.
 DR EMBL; S41242; AAB22680.1; JOINED.
 DR EMBL; S41243; AAB22680.1; JOINED.
 DR EMBL; S41244; AAB22680.1; JOINED.
 DR EMBL; S41245; AAB22680.1; JOINED.
 DR EMBL; S41246; AAB22680.1; JOINED.
 DR EMBL; S41247; AAB22680.1; JOINED.
 DR EMBL; S41248

DR EMBL: S40787; AAB22684.2; JOINED.
 DR EMBL: S40903; AAB22684.2; JOINED.
 DR EMBL: S40904; AAB22684.2; JOINED.
 DR EMBL: S40905; AAB22684.2; JOINED.
 DR EMBL: S40909; AAB22684.2; JOINED.
 DR EMBL: S40918; AAB22684.2; JOINED.
 DR EMBL: S40920; AAB22684.2; JOINED.
 DR EMBL: S40795; AAB22684.2; JOINED.
 DR EMBL: S40798; AAB22684.2; JOINED.
 DR EMBL: S40798; AAB22684.2; JOINED.
 DR EMBL: M68917; AAA41529.1; JOINED.
 DR EMBL: M68917; AAA41529.1; JOINED.
 DR EMBL: M68919; AAA41529.1; JOINED.
 DR EMBL: M68920; AAA41529.1; JOINED.
 DR EMBL: M68921; AAA41529.1; JOINED.
 DR EMBL: M68922; AAA41529.1; JOINED.
 DR EMBL: M68923; AAA41529.1; JOINED.
 DR EMBL: M68926; AAA41529.1; JOINED.
 DR EMBL: M68927; AAA41529.1; JOINED.
 DR PIR: A32460; A32460.
 DR PIR: A41343; A41343.
 DR HSSP: P22888; 1LUT.
 DR GCRdb: GCR_0138; -.
 DR GCRdb: GCR_0139; -.
 DR GCRdb: GCR_0262; -.
 DR GCRdb: GCR_0612; -.
 DR GCRdb: GCR_0613; -.
 DR GCRdb: GCR_0614; -.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 2.
 DR PRINTS: PR00373; GLYCHROMONER.
 DR PRINTS: PR01144; LSHRECEPTOR.
 DR SMART: SM00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 700
 FT DOMAIN 27 362 LUTROPIN-CHORIOGONADOTROPIC HORMONE
 FT TRANSSEM 363 390 RECEPTOR.
 FT DOMAIN 391 399 1 (POTENTIAL).
 FT TRANSSEM 400 422 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 423 443 2 (POTENTIAL).
 FT TRANSSEM 444 466 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 467 486 3 (POTENTIAL).
 FT TRANSSEM 487 509 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 510 529 4 (POTENTIAL).
 FT TRANSSEM 530 551 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 552 574 5 (POTENTIAL).
 FT TRANSSEM 575 598 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 599 609 6 (POTENTIAL).
 FT TRANSSEM 610 631 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 632 700 7 (POTENTIAL).
 FT REPEAT 52 75 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 126 150 LRR 1.
 FT REPEAT 152 175 LRR 2.
 FT REPEAT 176 200 LRR 3.
 FT REPEAT 202 224 LRR 4.
 FT REPEAT 225 248 LRR 5.
 FT REPEAT 250 271 LRR 6.
 FT DISULFID 443 518 LRR 7.
 BY SIMILARITY.

Query Match 9.4%; Score 474; DB 1; Length 700;
 Best Local Similarity 21.5%; Pred. No. 4.5e-25;

Matches 203; Conservative 130; Mismatches 322; Indels 288; Gaps 30;
 QY 5 PGIIRALMLCAAL-----CASRRAGAPGPPTACPAQCQEDGIMLSADCSSEIGLS 58

Db 6 PALRQLLVAVLVLLKPSQLOSRELG-----SRPECCDAPGAL----- 46
 QY 59 AVGDLDPLTAVLDLSMNNLTLEQPGFLHRLFELERLSGNLSHPGQAFGLYS-LK 117
 Db 47 RCPG---PRAGLARLSLTLP-----VKYIPSAFGLNEVYK 81
 QY 118 IMLQNNQGGIPAEALWELPSIQSLRLD--ANLISLVERSEGLSSRLHMLDQNAL 175
 Db 82 IEISQDSLSERLEANAEDNLNLSELLIONTKMLITYEP-GAFTNLPRLKYLISICWNGR 140
 QY 176 EIP-VALNNLPALQMTALN-RISHIDYAFARONLS-LVYLHNNRNQHLGTSFEG 232
 Db 141 TLPDVKIKISSSEFNFLTEICDNLHITTPQNAFQGMNNSVTLKLGNEFEVQSHAFNG 200
 QY 233 LNLLETLDLNYKLOEPVPAIRTLGRLOELGFHNNNIKAIPKAFMGNDLQTIHFYDNP 292
 Db 201 T-TLISLEKEN----- 211
 QY 293 IQFVGSARFOYLPKLTLSLNGAMDIOEPDLKGTTSLEILTLTRAGITLDSGMCQQLP 352
 Db 212 -----TYLEKMSGAFQATG-----PS----- 229
 QY 353 RLRLVELSHNQLEELPSLRQCQLEIEIGLHNRIMEIGADTFQSQSLQALDSMNAIRS 412
 Db 230 ---ILDISTKIQALPS-----HG-----LESIQTL----- 252
 QY 413 IHPPEAFSTLHSLVKLDLDNDQLTLPPLAGLGIMHLKIKGNLALSOAFSKDSEPKRLILE 472
 Db 253 -----IALSSYSKLTLP-----SKKFTSLIVAT 276
 QY 473 VPYAYOCCPYGWCASFEFRASGQWEADHLDESSKRPLGLLARAQENH--YDDDLDEL 530
 Db 277 LTYPSHCCAFRLPK-----KQONSFSITFENFSKQCESTVKRADETLYSAIFEN 328
 QY 531 QIEMED-----SKHPVSQCSPTPGPFPCCEYLFEFSWGRILAVMAIVLVSLCNGVLIT 585
 Db 329 ELSCWDYDYGFCSP-KTLQCAPEPDAPNFCEDIMGYAFRLVIMLINITAIFGN-LTYLF 386
 QY 586 VFAGSAPLPPVKPFVVGAINAGNTLTIGSCGLASVDALTIQOSEYGARWETGLGCRAT 645
 Db 387 VLLSRVLTVPFRLCMNLSPADFQCMGLYLLILASVDQTKQOYVNHADIMQGTGSCGAA 446
 QY 646 GFLAVLGSEASVLLTLTAAVOCVSVCRAVYAGKSPSGSVRAGVLGLALAGLAALPL 705
 Db 447 GFETVFASELVYTLFTTLERWHITTYAVQDDQRLRHAIPIMLGWLFSTLIATMPL 506
 QY 706 ASVGEYGAAPLCLPYAPPEGQPALGFTVALVMNSFCFLVAGAYIKLYCDLPRGFEA 765
 Db 507 VGISNYMKVASICLPM--DVESTLSQVYLLITILINVAFAVYICACYIRIYAVQNPDELTA 564
 QY 766 V-WCCAMVRYHVAWLIIPADGLLYCPVAFLSPASMLGFLPYTEAKSVLVLPACINP 824
 Db 565 PNKTKTIKKAAIILFTDFTCMAPISSFAAKRVPLITYNSKILLVLYEPVNSCANP 624
 QY 825 LLYLLENPHFRDRLRLRPRAAGSGLAYAAAGLEKSSCDSTQALYAFSDVDLLEASE 884
 Db 625 FLVAILFTKAFQDRFLLLISRFG-----CKKRA-----E 653
 QY 885 AGRPGLETY-----GFP-----SVTLSCQOPGAPR 911
 Db 654 LYRKKEFSAYTSNCKNGPQASKPSQATLKLSTVHCQDPIPR 696

RESULT 14
 TSHR_SHEEP
 ID TSHR_SHEEP STANDARD; PRT; 764 AA.
 AC P56455;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thyrotropin receptor precursor (TSH-R) (thyroid stimulating hormone
 receptor).

OS	TSHR .	Ovis aries (Sheep).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC	Bovidae; Caprinae; Ovis.	
OX	NCBI_TaxID=9940;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Pituitary;	
RX	MEDLINE=97445147; PubMed=9299474;	
RA	Boekmann J., Witter C., Witkowski W., Kreutz M.R., Boeckers T.M.,	
RT	"Cloning and expression of a brain-derived TSH receptor."	
RL	Biochem. Biophys. Res. Commun. 238:173-178(1997).	
CC	-1- FUNCTION: RECEPTOR FOR THYROID HORMONE. PLAYS A CENTRAL ROLE IN	
CC	CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS	
CC	RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE	
CC	CYCLASE.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.	
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.	
CC	FSH/LSH/TSH SUBFAMILY.	
CC	-1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/sib-ch).	
CC	or send an email to license@sib-sib.ch).	
CC	-----	
DR	EMBL: Y13434; CAA73846.1; -.	
DR	HSSP: P16473; 1XUM.	
DR	GCRCDB: GCR_2587; -.	
DR	InterPro: IPR000276; GPCR_Rhodopsn.	
DR	InterPro: IPR001611; LRR.	
DR	Pfam: PF00001; 7tm_1; 1.	
DR	Pfam: PF00560; LRR: 3.	
DR	PRINTS: PR00373; GLYCHROMONER.	
DR	PRINTS: PR01145; TSHRECEPTOR.	
DR	PROSITE: PS00237; G-PROTEIN_RECPEP_F1_1; 1.	
DR	PROSITE: PS50262; G-PROTEIN_RECPEP_F1_2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;	
KW	Phosphorylation; Repeat; Leucine-rich repeat.	
FT	SIGNAL	1 21
FT	CHAIN	22 764
FT	DOMAIN	22 413
FT	TRANSMEM	414 441
FT	DOMAIN	442 450
FT	TRANSMEM	451 473
FT	DOMAIN	474 494
FT	DOMAIN	495 517
FT	DOMAIN	518 537
FT	TRANSMEM	538 560
FT	DOMAIN	561 580
FT	TRANSMEM	581 602
FT	DOMAIN	603 625
FT	TRANSMEM	626 649
FT	DOMAIN	650 660
FT	TRANSMEM	661 682
FT	DOMAIN	683 764
FT	DISULFID	494 569
FT	REPEAT	51 74
FT	REPEAT	125 150
FT	REPEAT	152 174
FT	REPEAT	176 199
FT	REPEAT	201 223
FT	REPEAT	225 248
FT	CARBOHYD	77 77
FT	CARBOHYD	99 99
FT	CARBOHYD	177 177
FT	CARBOHYD	198 198
FT	CARBOHYD	302 302
SEQUENCE	764 AA; 86674 MW; 32423497056EB896 CRC64;	

Query Match	Similarity	9.3%	Score 473.5	DB 1	Length 764
Best Local	Similarity 23.4%	Pred. No. 5	5e-25		
Matches	206	Conservative	120	Mismatches	322
				Indels	23
				Gaps	35

DE Lutropin-Chorionadotropic hormone receptor precursor (LH/CG-R)
 DE (LSH-R) (luteinizing hormone receptor).
 GN LHCR OR LHR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92165799; PubMed=1311310;
 RA Gudermann T., Birnbaumer M., Birnbaumer L.;
 RT "Evidence for dual coupling of the murine luteinizing hormone
 RT receptor to adenylyl cyclase and phosphoinositide breakdown and Ca²⁺
 RT mobilization. Studies with the cloned murine luteinizing hormone
 RT receptor expressed in L cells.";
 RL J. Biol. Chem. 267:4479-4488(1992).
 RN [2]
 RP SEQUENCE OF 1-58 FROM N.A.
 RX MEDLINE=93093308; PubMed=1459341;
 RA Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;
 RT "The murine luteinizing hormone and follicle-stimulating hormone
 RT receptor genes: transcription initiation sites, putative promoter
 RT sequences and promoter activity.";
 RL Mol. Cell. Endocrinol. 88:55-66(1992).
 CC -1- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIONADOTROPIC HORMONE.
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
 CC ACTIVATE ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S49753; AAB24402.1; -;
 DR EMBL: M81310; AAA39432.1; -;
 DR EMBL: M87571; AAA39433.1; -;
 DR PIR: A42395; A42395.
 DR HSSP: P22888; 1LUT.
 DR GCRDB: GCR 0189; -;
 DR GCRDB: GCR 0305; -;
 DR MCD: MGI:96783; lhcgr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7em_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR PRINTS: PR00373; GLYCHROMONER.
 DR PRINTS: PRO1144; LSHRECEPTOR.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 26
 FT CHAIN 27 700
 FT DOMAIN 27 362
 FT TRANSSEM 363 390
 FT DOMAIN 391 399
 FT TRANSSEM 400 422
 FT DOMAIN 423 443
 FT TRANSSEM 444 466
 FT DOMAIN 467 486
 FT TRANSSEM 487 509
 FT DOMAIN 510 529
 FT TRANSSEM 530 551
 FT SIGNAL 551 551
 FT TRANSSEM 551 551

FT DOMAIN 552 574 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 575 598 6 (POTENTIAL).
 FT DOMAIN 599 609 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 610 631 7 (POTENTIAL).
 FT DOMAIN 632 700 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 52 75 LRR 1.
 FT REPEAT 126 150 LRR 2.
 FT REPEAT 176 200 LRR 3.
 FT REPEAT 225 248 LRR 4.
 FT DISULFID 443 518 BY SIMILARITY.
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 700 AA; 78214 MW; 8A6840A011E1E014 CRC64;
 Query Match 9.3%; Score 472; DB 1; Length 700;
 Best Local Similarity 21.3%; Pred. No. 6.2e-25;
 Matches 197; Conservative 147; Mismatches 331; Indels 250; Gaps 31;
 QY 5 PGLRALMICALCARRAGAPQPGPRACAPCHQEDGIMLSADCELSAVPGCL 64
 D 6 PALQLLVLAIVLKQSLHSP-LSGRCPEDCDAGAL-----RCGP-- 50
 QY 65 DPLAVLDLNNMNLTELQGLFHHRLFELRLSGNHSHINGQAFSGYS-LKIIMLGN 123
 D 51 -PRAGLALSLTYLP-----YKVIQSQAFRGINEVVKIISQS 87
 QY 124 NQLGIPAEALMELPLOSRLDANLISLPERSEFSGSLRLHMLDANA-LTEIPVRAL 182
 D 88 DSL------RIEAN-----AFDNLNLSEILIONKNLTYIEPGAF 123
 QY 183 NNPLALQMTALNRISHIIPYAFONLTS-VYLHNN-RQHGTHSEFEGHNETFD 240
 D 124 TNLPLKLTLSNGMIRLPDVSKISSEFNFLICDLYTTTPGAFQGMNN-ESIT 182
 QY 241 LNVNKLQEPVAIRTLGRLOELGFHNNNIKAIPKAFMGNPLQTLHFQDNQLQFGRSA 300
 D 183 LK-----LYGNFEEVQSNAFNQTTIL-SLELKEN----- 211
 QY 301 FOYLPKLTLSLNGAMDIOEPFDLKGTTSLTLLTLRAGIRLLPSGMCQDLRLVLELS 360
 D 212 -IYLEKMSGTFGATG-----PS-----ILDVS 234
 QY 361 HNOIEELPSLRCKLEIEIGLOHNRIMEIGADTFESQSLQALDLSWNAIRSIHPAEST 420
 D 235 STKLQALPS-----HG-----LESIQTL-----IATSSYS- 259
 QY 421 LHSIVKLDLTDNQLTTLPLAGIGLMHLKLGNALLSQAFSKDSEPKLILEVPAVQCC 480
 D 260 -----LKTLP-----SKEKTSLLVALITLYPSHCC 284
 QY 481 PYGMCASFEEKASQWEADLHLDESSKRPGLLAROAEHN-YDQDDELQ-----E 533
 D 285 AFRNLPR-----KEQNFSPSIFENFSKQCESTVRANNETLVSAPREEMELSGWDY 336
 QY 534 MEDSKPHSVQCCSPRPGFKCEYLFESWGIPLAWAIVLVSQNGVLLVLPFAGRP 593
 D 337 YDFCSP-KTLOCTEPDPAFNPCEDEIMGYAFLLVLLINILAIIFGN-LTVLFLVLTSSRYK 394
 QY 594 LPYKFFVAGLAGANTLTGICGLASVADALFGQFSEYGARWETGLCRATGFLAVLGS 653
 D 395 LTVRFELMCNLSFADFCGLVILLIASVDSQTKGYYNHAIDMOTGSGCSAAGFTTVRAS 454
 QY 654 EASVILLTLAAVQSVSVYCARVKGSPSLGVSRAVGLCLALAGIAALPLASVGEYGA 713
 D 455 ELASYITLVITLERHNTITVAVDQKRLRNRIAPIMAGWFTSLMTPLPLVGVSSYVK 514
 QY 714 SPLCLPAAPRPGQPAAGFTVALVMMNSFCPLVNAVAYIKLYCDL-PRQDFEAV-WDCAMV 772
 D 515 VSICLPM--DVESTLSQYIILILLNNAVAVVICACVRIYFAVQNELTAPNRDQTKIA 572

```
QY 773 RHVAMLIFFADGLLYCPVAFLESPASMLGIFPVTPEAVKSVLLVLPUPACUNPLLYLFNP 832
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 573 KKMALIFTDFTCMAPISFFAISAAEKVPLITYNSKVLVLYFPVNSCANPFLYAVFTK 632
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
QY 833 HFRDDLRLRPRAAGDSGLPLAYAAGLEKSSCDSTQALVAFSVDVLTILEASEAGRPPGLE 892
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 633 AFQRDFLLLSRFG-----CCKHRAELYRRKEFSACTFNKNGFPRSSK 676
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
QY 893 TYGFPS-----VTLISCQOPGAPRL 912
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 677 ----PSQALKLSIVHCQOPTPRV 697
   : : | | | | : : | : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: October 23, 2002, 10:42:49
Job time : 26 secs

✓

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52 ; Search time 42 Seconds
(without alignments)
3983.002 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MESPPELRALMICALCAKSR.....GGLSGGCGPQPSGLAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organellar:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4842	95.5	928	4	Q9BYD7 homo sapien
2	4092	80.7	828	4	Q9HBX8 homo sapien
3	2525.5	49.8	907	4	Q75473 homo sapien
4	2514.5	49.6	907	4	Q9UP75 homo sapien
5	2464.5	48.6	907	11	Q9Z1P4 mus musculus
6	2023.5	40.0	951	11	Q9Z2H4 ratu
7	2023.5	39.9	951	4	Q9BXB1
8	2012.5	39.7	951	4	Q9BYD1
9	1206	23.8	230	4	Q96K69
10	1095.5	21.6	1360	5	Q9ND11
11	1017.5	20.1	1012	5	Q95Y16
12	1017.5	20.1	1280	5	Q95Y17
13	1010.5	19.9	1300	5	Q9NKG6
14	941	18.6	1050	5	Q9BN18
15	546	10.8	701	13	Q9DGC6
16	514.5	10.2	724	13	Q9PYV0

17	512.5	10.1	688	11	Q64183
18	510.5	10.1	693	13	Q9DGC5
19	503	9.9	779	13	Q91BN7
20	491.5	9.7	829	5	Q9VEG4
21	490.5	9.7	831	5	Q94979
22	486.5	9.6	658	13	Q9PVN9
23	486.5	9.6	778	13	Q98TP4
24	486	9.6	764	11	Q9D697
25	484.5	9.6	1091	11	P70193
26	477	9.4	763	6	Q9BGN4
27	473	9.3	1093	4	Q96JAL
28	472.5	9.3	793	13	Q91948
29	472	9.3	814	13	Q91949
30	472	9.3	737	11	Q9HYB8
31	463	9.1	764	6	Q9BGS6
32	462.5	9.1	603	11	Q70211
33	459.5	9.1	603	13	Q9DGE5
34	459	8.9	739	6	Q9BGS5
35	452	8.9	662	13	Q9BT84
36	447	8.8	662	13	Q9PWL6
37	445	8.8	757	4	Q9HBX9
38	442.5	8.7	603	11	Q9PB17
39	439.5	8.7	699	4	Q15996
40	439.5	8.7	687	11	Q9U1L0
41	438.5	8.7	696	13	Q9BT85
42	435.5	8.6	696	13	Q9BT85
43	434.5	8.6	570	11	Q9DBB9
44	433.5	8.6	601	13	Q42500
45	432.5	8.5			

ALIGNMENTS

RESULT 1

Q9BYD7 PRELIMINARY: PRT: 928 AA.
AC Q9BYD7;
DT 01-JUN-2001 (TREMUREL, 17, Created)
DT 01-JUN-2001 (TREMUREL, 17, Last sequence update)
DT 01-DEC-2001 (TREMUREL, 19, Last annotation update)
DE VTS20631 (FRAGMENT).
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki H., Hayashi A., Kozuma S., Saito T.;
RT "a member of g-protein coupled receptor family."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB049405; BAB39854.1; --
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_in.
DR Pfam: PF00560; LRR_15.
DR PRINTS: PR00237; GPCRHDOPSN.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR_9.
DR SMART: SM00369; LRR_TYP; 14.
FT NON_TER
SQ SEQUENCE 928 AA; 100487 MW; 4C3364AD8A9C463 CRC64;

Query Match 95.5%; Score 4842; DB 4; Length 928;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 HCOEDGIMLSADSCSELGSAVPGDLPFLYALDLSNNLTTELPGFLFHRLFEELRLSG 99
DB 1 HCOEDGIMLSADSCSELGSAVPGDLPFLYALDLSNNLTTELPGFLFHRLFEELRLSG 60

```

QY 100 NMLSHIPGQAFSGLYSLKILMLONNOLGIPAPALMELPSLOSILRUDANLSLVERSE 159
DB 61 NMLSHIPGQAFSGLYSLKILMLONNOLGIPAPALMELPSLOSILRUDANLSLVERSE 120
QY 160 GISSLRHMLDNLALTEIPVRLANMLPALQAMTALNIRISHIPDYAFONTLSVLVLIHN 219
DB 121 GISSLRHMLDNLALTEIPVRLANMLPALQAMTALNIRISHIPDYAFONTLSVLVLIHN 180
QY 220 NRIQHLGTHSEFGLNLETLDTLNKLOEPVAPITTLGRLOELGFHNNNKAIPERKAFMG 279
DB 181 NRIQHLGTHSEFGLNLETLDTLNKLOEPVAPITTLGRLOELGFHNNNKAIPERKAFMG 240
QY 280 NPLIOTIFYTNPLOFVGRSAFOYLPRKLTLSLNGAMDIOEPDLKGTSTSEITLIRAG 339
DB 241 NPLIOTIFYTNPLOFVGRSAFOYLPRKLTLSLNGAMDIOEPDLKGTSTSEITLIRAG 300
QY 340 IRLPSCMCQOLPRRLVLELISHNOIEELPSLRQKLEELGLOHNRIMEIGADTFSSQSS 399
DB 301 IRLPSCMCQOLPRRLVLELISHNOIEELPSLRQKLEELGLOHNRIMEIGADTFSSQSS 360
QY 400 LQALDLSWMAIRSHIHPAFSTLSLVKLDITDNLQTLPLAGLGGMLHLKGNLALSOA 459
DB 361 LQALDLSWMAIRSHIHPAFSTLSLVKLDITDNLQTLPLAGLGGMLHLKGNLALSOA 420
QY 460 FSKDSFKRLILEVPAVYOCCTPGKCSAFKASQWAEEDLHDDRESSKRPGLIARQA 519
DB 421 FSKDSFKRLILEVPAVYOCCTPGKCSAFKASQWAEEDLHDDRESSKRPGLIARQA 480
QY 520 ENHYDDDLDELQLEMEDSKPSPVQCSPPPGPKPCEYLFEESGILAWAYILSVLCN 579
DB 481 ENHYDDDLDELQLEMEDSKPSPVQCSPPPGPKPCEYLFEESGILAWAYILSVLCN 540
QY 580 GVLILVFEAGGAPRLPRPVFVVGALAGANTLQISGCLLASVDALTFQGFSEYGARWETG 639
DB 541 GVLILVFEAGGAPRLPRPVFVVGALAGANTLQISGCLLASVDALTFQGFSEYGARWETG 600
QY 640 LGCRAATGFLAVIGSEASVLLTLAAVQCSVSVSCVRAVAKSSISGSVRAVYIGCLALAGL 699
DB 601 LGCRAATGFLAVIGSEASVLLTLAAVQCSVSVSCVRAVAKSSISGSVRAVYIGCLALAGL 660
QY 700 AALALPLASGEYGAFLCLPYAPPEGQAPALGFTVALVMNSFCFLVVAAGATIKLYCDLP 759
DB 661 AALALPLASGEYGAFLCLPYAPPEGQAPALGFTVALVMNSFCFLVVAAGATIKLYCDLP 720
QY 760 RGDFAVMDCAVRRVAVMLIFADGILLYCPVAFSLFASMLGFPTYPEAVKSVLLVPLP 819
DB 721 RGDFAVMDCAVRRVAVMLIFADGILLYCPVAFSLFASMLGFPTYPEAVKSVLLVPLP 780
QY 820 ACLANLILLYLNPFRDLRLRPRAGDSGLPAAAGELKSSGCDSTQALVAFSDVDLI 879
DB 781 ACLANLILLYLNPFRDLRLRPRAGDSGLPAAAGELKSSGCDSTQALVAFSDVDLI 840
QY 880 LEASBAGRPGLIETGFPVSYTLISQOGAPRLGSGSHRVEPEGNHFGNPOPSMDGELLR 939
DB 841 LEASBAGRPGLIETGFPVSYTLISQOGAPRLGSGSHRVEPEGNHFGNPOPSMDGELLR 900
QY 940 AEGSTPAGGGLSGGGGPOPSGLAFASHY 967
DB 901 AEGSTPAGGGLSGGGGPOPSGLAFASHY 928

```

RESULT 2

Q9HBX8

ID Q9HBX8 PRELIMINARY; PRT; 828 AA.

AC Q9HBX8; 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DE LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-COUPLED RECEPTOR 6

DE (FRAGMENT).

GN LGR6

OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; Pubmed=10935549;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RT van der Spek P.J., van Duin M., Hsueh A.J.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signaling mechanism for LGR7."
RL Mol. Endocrinol. 14:1257-1271(2000).
DR EMBL; AF190501; AAG17168.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 10.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEUCRICHPT.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00369; LRR_type; 10.
KW Receptor.
FT NON_TER
SQ
SEQUENCE 828 AA: 89387 MW: 1209251EA442DB4 CRC64;
Query Match
Best Local Similarity 80.7%; Score 4092; DB 4; Length 828;
Matches 794; Conservative 89.1%; Pred. No. 5.2e-301;
Indels 96; Gaps 1;
QY 72 DLSNNLELQPGFLFHHRLFLBELRLSGNHLSHIPGQAFSGLYSLKILMLONNOLGIPA 131
DB 29 DLSNNLELQPGFLFHHRLFLBELRLSGNHLSHIPGQAFSGLYSLKILMLONNOLGIPA 88
QY 132 EALMELPSLOSILRUDANLSLVERSEFGLSLHMLDNLALTEIPVRLANMLPALQAM 191
DB 89 EALMELPSLOSILRUDANLSLVERSEFGLSLHMLDNLALTEIPVRLANMLPALQAM 99
QY 192 TLALNRISHIPDYAFONTLSVLVLIHNRIOLGTHSEFGLNLETLDTLNKLOEFPV 251
DB 100 -----LDLNKLOEFPV 112
QY 252 AIRTLGLOELGFHNNNKAIPERKAFMGNPLQTIHYDNPLOFVGRSAFOYLPRKLTLS 311
DB 113 AIRTLGLOELGFHNNNKAIPERKAFMGNPLQTIHYDNPLOFVGRSAFOYLPRKLTLS 172
QY 312 LNGAMDIOEPPDLKGTSTSEITLTFRAGRILPLPSGMCQOLPRRLVLELISHNOIEELPSLH 371
DB 173 LNGAMDIOEPPDLKGTSTSEITLTFRAGRILPLPSGMCQOLPRRLVLELISHNOIEELPSLH 232
QY 372 RCKLEELIGLOHNRIMEIGADTFSSQSSLOALDLSWMAIRSHIHPAFSTLSLVKLDITD 431
DB 233 RCKLEELIGLOHNRIMEIGADTFSSQSSLOALDLSWMAIRSHIHPAFSTLSLVKLDITD 292
QY 432 NQITTLPLAGLGLMLHLKGNLALSOAFSKDSFPLRILEVPAYVOCCTPGKCSAFKFA 491
DB 293 NQITTLPLAGLGLMLHLKGNLALSOAFSKDSFPLRILEVPAYVOCCTPGKCSAFKFA 352
QY 492 SGQWEAEDLHDDRESSKRPGLIARQANHYDQDLDELQLEMEDSKPSPVQCSPPG 551
DB 353 SGQWEAEDLHDDRESSKRPGLIARQANHYDQDLDELQLEMEDSKPSPVQCSPPG 412
QY 552 FKPCYELFESWIGIRLAWAYILSVLCNGVILYTFAGGAPRLPRPVKRVGALIGANTLT 611
DB 413 FKPCYELFESWIGIRLAWAYILSVLCNGVILYTFAGGAPRLPRPVKRVGALIGANTLT 472
QY 612 GISCGLLSVALTFFGQSEYGARWETGLGCRATGFLAVIGSEASVLLTLAAVQCSVSV 671
DB 473 GISCGLLSVALTFFGQSEYGARWETGLGCRATGFLAVIGSEASVLLTLAAVQCSVSV 532
QY 672 SCVRAVYKSPSLGVRAGVGLCLALAGLAALPLASGEYGAFLCLPYAPPEGQAPALG 731
DB 533 SCVRAVYKSPSLGVRAGVGLCLALAGLAALPLASGEYGAFLCLPYAPPEGQAPALG 592

```

```

RESULT 4
Q9UP75
ID Q9UP75 PRELIMINARY; PRT: 907 AA.
AC Q9UP75;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE G PROTEIN-COUPLED RECEPTOR LGR5.
GN LGR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.G., Hsueh A.U.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a
RT * protein-coupled, seven-transmembrane region."
RL Mol. Endocrinol. 12:1830-1845(1998).
DR EMBL; AF061444; AAC7911.1; -.
DR HSSP; P23945; IXUN.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.

```



```

RESULT 7
Q9BXBI PRELIMINARY: PRT: 951 AA.
AC Q9BXBI.
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE G-PROTEIN COUPLED RECEPTOR.
GN GPR48.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21294803; PubMed=11401528;
RA Loh E.D., Brousard S.R., Kolakowski L.F.;
RT "Molecular Characterization of a Novel Glycoprotein Hormone G-Protein-
RT Coupled Receptor."
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
DR EMBL; AF346711; AAK31153.1; -.
DR EMBL; AF346709; AAK31153.1; JOINED.
DR EMBL; AF346710; AAK31153.1; JOINED.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Mterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000822; znf-C2H2.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR.15.
DR Pfam; PF01462; LRRNT.1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00369; LRRNT.1.
DR SMART; SM00369; LRR_Typ.15.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Receptor.
SQ
SEQUENCE 951 AA: 104460 MW: 56032DFCF22CA1BB CNG64;

Query Match 39.9%; Score 2023.5; DB 4; Length 951;
Best Local Similarity 46.3%; Pred. No. 2.4e-144;
Matches 432; Conservative 130; Mismatches 303; Indels 69; Gaps 12;

QY 1 MSPRGLALMICALCSRRAGARQREGPTACPARCHQEDGIMLSADCSGLISAY 60
DB 1 MGPPLGL---LC--FLALGLLSAGSPGGAAPPLCAAPCSGCDG--RRVDCSGKGLTAV 51
QY 61 PGDDPLAVYLDLSNNNTLELPGLFHLRLFEELRLSGNHLSHIPGQAFSLYLKILM 120
DB 52 PEGLSAFQALDISNNNTIQLPEDAFKNPFLLEQLQLAGNDLSFTHPKLSGLKELKYL 111
QY 121 LONNOLGIRPEALWELPSLOSRLDANLISLVPERFEGSLSLRHLMDNATEIPVR 180
DB 112 LONNOLGIRPEALWELPSLOSRLDANLISLVPERFEGSLSLRHLMDNATEIPVR 171
QY 181 ALNNLPALQATLALNRISHIPDAFQNLSTLVYLHLNNNTLOHLGTHSFGSLNLEFLD 240
DB 172 PLNNLPALQATLALNRISHIPDAFQNLSTLVYLHLNNNTLOHLGTHSFGSLNLEFLD 231
QY 241 LNYKKLOEFPVAITLQLOELGFGHNNNIKAIPKAFGPNLQTIHYNDPIQFVGS 300
DB 232 LSYNNLGEFPVAIKARSLKELGFGHNSISVIPGADGPNLQTIHYNDPIQFVGS 291
QY 301 FOYLPKHTLSLNGAMDIQEPDLKGTTSLEILTLPFAGIRLLPSGMCQOLPRLVLELS 360
DB 292 FHNLSLHSLVIRKASVVOQPNLTGTVHLESLTLTGTKISSIPNNLCQKQMLRTDLS 351
QY 361 HNOIEELPSLRCKLEIELOHNRIVEIGADFPQSLSQALDLSMNAIRSHPEAFST 420
DB 352 YNNIRDLPSNGCHALEISLOHNOITQIEKGTFQGLISRLIDLSKNLILHEIHSRAFAF 411

```

```

QY 421 LHSIVKLDITDNOITLPLAGIGLMHLKGNLALSOAFKSDSPFKLITLPEVAYQCC 480
DB 412 LGPITNLDVSEFNEILTSFPEGLNGINGLOKLVGNFKLKELAAKDFVNRSLSTVPAVQCC 471
QY 481 PYGMCASFEEKASGOWEAEDLHDD----EESKRLPGLILARAENHYDODDELQLEMD 536
DB 472 AFWGDSY-----ANLTDENSLQDHSVAOEKGTADANANTSTLEENHSQI----- 518
QY 537 SKPHSVQCSPPPEPFKCEYLFESWGIRLAWYAVLTVSLCNGILVLTVPAGPAPPP 596
DB 519 -----IHCPTSGAFKCEYLLGSWMIRLTVWFIFVLALFENLTVLTTFASCS -LPS 572
QY 597 VKFVGAAGANTLTLGICSGILASVDALTFQGFSEYKARWETGLGCRATGLAVIGSEAS 656
DB 573 SKLFIGLSTVSNLPMGITTYTGLITFLDAVSKGRFAEGTWEFTSGCKKAGFLAVFSSESA 632
QY 657 VLLTLAAVQCSVSVSCVRAVYKSPSLGSRVAGVLCIALAGLAAALPLASVGEYASPL 716
DB 633 IFLLMLATVERSLSAKDIMKKNHLLKQFVAAALLAFLGATVAGCPLFTFRGEYSASPL 692
QY 717 CLPAPPEGQPAALGFYVATVMNMSFCFLVYAGAYIKYCDLPBGDEAVAMCAMVHVA 776
DB 693 CLPP--PTGETPSLGFVTYVLLNSLAFLLMAVITYTKLYCLEDELSSENSOSMIKHA 750
QY 777 WLTPADGLYCPVAFLPSASMLGLEPYTPKAVSVLTVPLPACLPBLTYLTPNPHRD 836
DB 751 WLTPNCIFFCPVAFSFAPLITNISPELMKSVTLIFPLPACLPBLTYLTPNPHRKE 810
QY 837 D---LRLRPRAGDSPLVYAAAGLEK-----SSCDSTQALVAFSDV 876
DB 811 DWKLKRRVTKKSGSVSVSSQGGCLEDPYDYGMYSHLQGLNLFVCDCESEFLTKPV 870
QY 877 D---LLEASAGRPPLLETYGFPSVTLISGQOP 907
DB 871 SKCHLI-----KSHSCPALAVASCORP 892

RESULT 8
Q9NYD1 PRELIMINARY: PRT: 951 AA.
AC Q9NYD1.
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE G-PROTEIN COUPLED RECEPTOR 48.
GN GPR48.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Loh E.D., Brousard S.R., Kolakowski L.F., Jr.;
RT "Molecular Characterization of A Novel Human Glycoprotein Hormone G-
RT Protein Coupled Receptor."
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257182; AAF68989.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Mterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000822; znf-C2H2.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR.15.
DR Pfam; PF01462; LRRNT.1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00013; LRRNT.1.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.

```

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 951 AA; 104358 MW; 88DA6C67F6ABA1FE CRC64;

Query Match 39.7%; Score 2012.5; DB 4; Length 951;
 Best Local Similarity 46.0%; Pred. No. 1,6e-143;
 Matches 430; Conservative 130; Mismatches 305; Indels 69; Gaps 12;

QY 1 MESPRLRLMILCAALCASRRAGGAPQPGPGPACAPRCQDQDGLMLSDSGELGSAV 60
 DB 1 MESPRLGL----LC--FLAGLGLSAGSPSGAPPLCAAPCSCDGD--RRYDCSGKGLTAV 51
 QY 61 PGDLDELTAVALDLSMNNLTLELQGLFHHRLFELELRLSGNHLSHIPQAFSGLYSTKIM 120
 DB 52 PEGLSAFQALDLSMNNITQLPEDAPKPNFLELQLAGNDLSIRHKALSGLEKELKVL 111
 QY 121 LQNNQAGGAPAEALMELPQLQSLRLDANLISLVERSEFEGISIRHLMDNALTEIPVR 180
 DB 112 LQNNQKTVSEAIRGLSALQSLRLDANHTSYDEDFEGLVQRLHMLMDNSLTEVPVH 171
 QY 181 ALNNRLQAMTLALNRISHIIPYAPQNLISLVYLHNNRIRGLGTHSFEGLHNETLD 240
 DB 172 PLNLPTQLQTLTALNKLSSIPDFAFTNLSLVVHLHNNKIRGLSGHCEFDGLDNETLD 231
 QY 241 LNNKLOEFPVALRTGLQELGFHNNNKAIPKAPMGNPLOTIFHYNDPIQFVGRSA 300
 DB 232 LSTNNLGEFOAIKARSLKELGFHSNISIVIPDAGDPDLRTIHLNDPLSFVGN 291
 QY 301 FOYLPKRLHTLSLNGAMDIOEPDLKGTSLLEITLRAGIRLLPSGMCQQLPRLVLELS 360
 DB 292 SHNLSDHSLVIRGASVWQFPNLTGVHLESLTGTKISSIPNNLCQEQKMLRTLDLS 351
 QY 361 HNOIEELPSLRQKLELQHLNRIWEIGADPESQSLQALDLSMNAIRSHIEPAST 420
 DB 352 YNNIRDLPSFNGCHALEISLQRMQYQIKEGTFQGLSKLIDLSKRLHEIHSRAVAT 411
 QY 421 LHSIVKLDLNDQTLTTPPLAGLGLMHLKGNLALSQAESKDSFPLRIILEVYAAQC 480
 DB 412 LGPTINDVGFENELTSPTEGPNGLNDKLGNFKLEALAKPVDNRLSYVYAAQC 471
 QY 481 PYGMCASFRRASGQWEADLHLD-----ESSKRPGLLARQAENHYDQDDELQLEMED 536
 DB 472 AFMCDSY----ANLNTEDNSLDHSAQEKGTADAANVTSTLENEHSQT----- 518
 QY 537 SKRHPVOCSTPGPKPCETLFESMGIRLAVMAIVLISVLCNGLVLTVAAGPAPRL 596
 DB 519 -----IHCPTSGAFKPCETLFGSMIRLWFLVLFVLFNLVLTVAASCTSLPS 572
 QY 597 VKFVGAAGANTLTGISCGLASVDALTFEQSEYGAQMETGLGRATGFLAVLGSSEAS 656
 DB 573 SKLFIGLISVSNLFMGITYGLTFLDAVSWGRFAEPFGIMWETSGCKVAGFLAVFSSSA 632
 QY 657 VLLITLAAVQCSVSCVRAVAKSPISGVRAGVGLCLALAGLAALPLASVGEYASPL 716
 DB 633 IFLLMLATVERLSAKDIMKNGKSNHKKQFRVALLSAFLGATVAGCPPLFRHGEYSASPL 692
 QY 717 CLPAPRPGQPAALGTFTVAVLMMNSFCFLVYAGVYIKYCDLPBGDEFAVADCMVHRVA 776
 DB 693 CLP--PTGEPSPGLFTVTLVLLNSLAFVLYTKYLCWLEEDLSSENSOSSITKIVA 750
 QY 777 WLIFADGLLYCPVAFISFASMLGFVYTPVAEAVSVLVLPLACINPLVLTLPNPRFD 836
 DB 751 WLITNCLIFCPVAFISFAPLITAISSPELMKSVTLIFPLPACINVLVLFVFNPKKE 810
 QY 837 D-----LRLRPRAGDSGLAVAAAGLEK-----SSCDSTQALVAFSDV 876
 DB 811 DMKLKRRVTKKSGSVSISISOGGCEODFYDDGMYSHLQGNLTVDCCESEFLTRPV 870
 QY 877 D---LILEASAGRPGLTEYGFPSVTLISCOOP 907
 DB 871 SKHLLT-----KSHSCPALAVASCORP 892

RESULT 9
 Q96K69

ID 096K69 PRELIMINARY; PRT; 230 AA.
 AC 096K69;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CDNA FLJ14471 FIS, CLONE MAMMAL1001030, WEAKLY SIMILAR TO
 DE LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR.
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;

RA Tsogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Makamatsu A., Nakamura Y., Nagahari K., Masuko Y., Oshima A.;

RT "MEDO human CDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027377; BAB55071.1; -

SQ SEQUENCE 230 AA; 24403 MW; 3BF762BA34A9D14F CRC64;

Query Match 23.8%; Score 1206; DB 4; Length 230;
 Best Local Similarity 99.6%; Pred. No. 2.3e-83;

Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 738 MANSFCEFLVAGAVIKLYCDLPRGDDEAVWDCAMVRYAWLIFADGLLYCPVAFISFASM 797
 DB 1 MANSFCEFLVAGAVIKLYCDLPRGDDEAVWDCAMVRYAWLIFADGLLYCPVAFISFASM 60

QY 798 LGLPVTPAEAVKSVLYVLPAPCLNPLLYLFPNPFDDLRRLRPRAGDSGLPATAAG 857
 DB 61 LGLPVTPAEAVKSVLYVLPAPCLNPLLYLFPNPFDDLRRLRPRAGDSGLPATAAG 120

QY 858 ELEKSSCDSTALVAFSVVDLILEASEAGRPGLTETGFPSTVTLISCOOPAPRLGSHC 917
 DB 121 ELEKSSCDSTALVAFSVVDLILEASEAGRPGLTETGFPSTVTLISCOOPAPRLGSHC 180

QY 918 VEPGNNHGNPQPSMDGELLRAEGSTPAGGGLSGGGGFGPQSLAFASHV 967
 DB 181 VEPGNNHGNPQPSMDGELLRAEGSTPAGGGLSGGGGFGPQSLAFASHV 230

RESULT 10
 Q9NDI1

ID 09NDI1 PRELIMINARY; PRT; 1360 AA.
 AC 09NDI1;

DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE GLYCOPROTEIN HORMONE RECEPTOR II.
 GN RK OR BG:DS00180.13 OR CG8930.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON S.; TISSUE=WHOLE ANIMAL;

RA MEDLINE=20359836; PubMed=10899142;
 RA Erikson K.K., Hauser F., Schott M., Pedersen K.-M., Soendergaard L.,
 RA Grimmelikhuizen C.J.P.;

RT "Molecular Cloning, Genomic Organization, Developmental Regulation,
 and a Knock-Out Mutant of a Novel Leu-Rich Repeats-Containing G
 protein-Coupled Receptor (DLGR-2) from Drosophila melanogaster.";

RL Genome Res. 10:924-938(2000).

DR EMBL; AF142343; AAF6608.1; -
 HSSP; Q57815; IDBY.

DR FlyBase; FBgn0003255; rK.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_1p.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSN.
 DR SMART; SM00370; LRR; 2.
 DR SMART; SM00369; LRR_Typ; 5.
 DR PROSITE; PS50262; G_PROTEIN_RECPT_FL2; 1.
 KM Receptor.
 SQ SEQUENCE 1360 AA; 150731 MW; 7D435155B4F6F612 CRC64;

Query Match 21.6%; Score 1095.5; DB 5; Length 1360;
 Best local similarity 30.0%; Pred. No. 7.1e-74;
 Matches 292; Conservative 157; Mismatches 396; Indels 127; Gaps 18;

QY 39 CHCOEDGIMLSADCELSGLSAVPGDLDPLTAYLDLSMNNLTLPGLFHLRLLEIQLS 98
 DB 152 CHCTGSEVLRLSCRGIGLAVPVNLPNEVVLDGNNNLTKLEANSFMAPRLLEDLTLS 211
 QY 99 GNHLSHIGQAFSGLSXKILMLQNNQGGTAEALMEPISQSLRLDANLISLVERSEF 158
 DB 212 DNSITMDPNAFYGLAKLRSLQNGCKLSPPSFOGLAOLTSIQNGNALVSLDGDCL 271
 QY 159 EGLSSIRHLMDNALTEIPVRLNNLPALQMTLNLRIISHIPDVAFQNLTSVLVLLH 218
 DB 272 GHLQKRLRLLEGNFYRIPTNALAGRLLEALNLSNLTITINDEDFRMPRLVLLIK 331
 QY 219 NNRIOHGLTSHFEGIANLEETLDNLNKLQEFPAIRITGLRQELGHNHNKAIPEKATW 278
 DB 332 RNOIMKISAGALKNTALTAVLELDNLISSEGLSKISQDELSTSRMLRMINDTELP 391
 QY 279 GNPRLQTHFDNPRIQFGRSAFOYLRPKLHTLSLNGAMDIOEPRLDKTSTLEILTTRA 338
 DB 392 RS--MQMDMPANPLSTISAGFRGMSKIRKLISDVRTLRSPLEACHALEILKLDRA 449
 QY 339 GIRLPSGMCQOLPRLVLELSHNOIEELPSLHRC-----Q 374
 DB 450 GIOEPANLCKQTPRLKLELKTNSLRPNLSSCRDLRLDLSNQIKQCKPPNGIK 509
 QY 375 KLEITGLOHNRIMEIGADTFQSOLSLQALDSWNAIRSIHPEAFSTLSLVKLDITDNOI 434
 DB 510 QLANDLLSYNRKIKALPDQAFQIPIQLDLEGENEISYIHKBAFSGFTALIEDLNINIF 569
 QY 435 TTLPLAGIGGMHLKIKGNLALSOAFSKDSPKALILEVRYAQC--PYGMCASFPA 491
 DB 570 PELPESGLRALHLKTKTNPKLREFPPPTPRIQTLILSYAHCCAFPLVMASSQKKT 629
 QY 492 S-----GWEAEDLHLDDESSK 509
 DB 630 SOVGEAVLPESDAEDMTLMNNSMNIWPMNHLNLSKOLGASHMDWETA-IPNENQLOI 688
 QY 510 RPLGLLRAQENHYDQIDE-----IQLENEDESKRHPVQCSPTPGPFKPC 555
 DB 689 QTGQOIAITSYEEFEEDHVDGPRATGYGFTGLFSGMSTEDQD--GSVOCLPMPGFLPC 747
 QY 556 EYLEESGIRLAWVAIVLSTLNGIVLLTFAGGAPRLPRYKFFVGAAGANTLIGTIC 615
 DB 748 ADLEDMTILRCGVAVVFLSLNGSTVFLLC--SRKMDVPRLFLVCNLAADFQWGYIL 806
 QY 616 GLLASVDALTFQOSEYGARWETGLGRATGFLAVLGEASVLLTLTAAVQCSVSVCVR 675
 DB 807 GILAIYDAATLGERMFALPWQMSVLCQSLFVLAVLSSELST--YTLAVITIERNYATH 864
 QY 676 AYKGSPELGSYRAVGLCL--ALAGLAADPLASVEYGASPLCLPYAPPEGOPALAGT 733
 DB 865 AIHLNKRSLTKQAGIYISVGVFWALIMALPDLGVSDYRKFAVCPFPETTG--PASTLV 923
 QY 734 VALVMNNSFCFLVAVAGYIKYCDLPDEFAVW--DCAAVRHNAVLIFADGLIXCPVA 790
 DB 924 ISLAFITGCAFLITLMGCTLAKYTAI-RG--SOAMVNTDSRIAKRMALLVTFDFLCWSPIT 980

QY 791 FLTSASMLGELPPVPEAKSVLLVLPRLPACINPLLYLLFNHFRDRLRLRPR----- 844
 DB 981 FFSITAFIQLDLSLEQKIFITVFVLPPLNCSNCPPLVIMTKQFKDCTVTKHSESRV 1040
 QY 845 AGDSGPLAAAGALEKSSCDSTQALVAFSDVDL--LIEASEAGRPGLLETGYGFSYTL 901
 DB 1041 VGGGGPGRGAVARTKRG-----DLPPPLPAAVAHPPGCR-----CLNM 1081
 QY 902 ISCOQPGAPRL 913
 DB 1082 LPSEMPNWHKME 1093

RESULT 11

ID Q95Y16 PRELIMINARY; PRT; 1012 AA.
 AC Q95Y16;
 DT 01-DEC-2001 (Tremblrel, 19, Created)
 DT 01-DEC-2001 (Tremblrel, 19, last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, last annotation update)
 DE GLYCOPROTEIN HORMONE RECEPTOR.
 GN APGPHR'.
 OS Asterina pectinifera (Starfish).
 OC Asterozoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Eukaryota; Metazoa; Valvataceae; Valvatida; Asterinidae; Asterina.
 OX NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "DNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061862; BAB8209.1; -.
 KW Receptor.
 SQ SEQUENCE 1012 AA; 112623 MW; 52A70E7A88C4E0A CRC64;

Query Match 20.1%; Score 1017.5; DB 5; Length 1012;
 Best local similarity 30.2%; Pred. No. 3.7e-68;
 Matches 269; Conservative 132; Mismatches 360; Indels 131; Gaps 18;

QY 9 ALWLCALC-----ASRRAGAPQPGPPTACPAHCQOEGIMLSADCELSAVPGD 63
 DB 12 SYVILLLLCGRVVISVSGDPTGLCPGTLC---CH---DG-GHNVCSTRNLTDVPA 64
 QY 64 LDPITAYLDLSMNNLTLPGLFHLRLLELRLSGNHLSHIPQAFGLYSKLTMLON 123
 DB 65 LIGITETLISFNNSIILPADAFRHLPRDLTILIGNLSTLDKNVFGRLNLDLNLKL 124
 QY 124 NOLGTPAEALW--ELPSLOSLRLDANLISLVERSEFSGLSLHLMMDNALTEIPVRA 181
 DB 125 NRFQOYFRKAFRRNDLNLRLHDSNMWIREVPADAFNLTALHNLNDHQLSVEPTAA 184
 QY 182 LNNLPALQAMTLNLRIISHIPDVAFQNLTSVLVHLNHNRIQHGLTSHFEGIANLETDL 241
 DB 185 LHLHLSNRILHLEHNSIPVYDPAFAENSHLIELLRHNKTIHLSAFAFALPMLMLE- 243
 QY 242 NYKKLOEFPVAILTGLRQELGHNHNKAIPEKAFGNPLQITIHFDNPDIQVGRSAF 301
 DB 244 -----FLGNSITSTAH-----TAF 257
 QY 302 QYLPKLTHTSLANGAMDIOEPRLDKTSTLELTTLTRAGIRLRLPSGMCQOLPRLVLELSH 361
 DB 258 RNPALRLNVILEYKKNLSVPPDLTGTSLEHLGIERCSLRITPNNFCDDNMGSLNLDHN 317
 QY 362 NQIEELPSLHRCQKLE-----EIGLOHNRIMEIGADTFESQL 397
 DB 318 NLEGGPSLSKSGSKVTLHGTNKLTSLEQPFSGLDIVDQILLEDNDISTIPADAOQL 377
 QY 398 SSLQALDLSNNAIRSIHPEAFSTLSLVKLDITDNOITLPLAGLGLMHLKGNLALS 457
 DB 378 SHLDITLSNNTIREIDSOAFACPTSLQYLDLSNNSPVPVLTAGLQMLKIRTYDNEQLE 437

QY 458 QAFSKDSEPKRLILEVPYAYOCPPYGMCAEFKASQGWAEEDLHDDDESSKRPGLGLAR 517
 DB 438 DEPPSELPTSTETATATAPYHCCEYIELAEYLS-----LADRPNISETTYWASG 488
 QY 518 QAENHYDDDELQLEMEDS-----KHPVSQCSPTPGPEKP 554
 DB 489 SVDPYNNMTFEDINSWSIDSIFFGSLISIGSPTYLSGNTSRLVPH-NISCRPKRGPEMP 547
 QY 555 CEYFESMGIRLAWAYVLLSVLCNGVLLTVFAGGPARLPVKYVGAIAAGANTLTGIS 614
 DB 548 CMDLFGSMPRLIGWMLFELALIGNAIVFVIYVSHTRKMDVP-REFLGNLAFADFLGY 606
 QY 615 CGLASVVALTFGQSEYKAWETGLGCRATGFLAVLGSEAVLLTLTAAVOCYSVSCV 674
 DB 607 LGFLAGVDTSTLGVPRKRGARQWLSAGCRLAGFLAVFSESESYTLSTLTERFAIKHA 666
 QY 675 RAYGKSPSLGSRVAGVIGCLA--LAGLAALPLASVGEYASPLCLPYAPPEGOPALGF 732
 DB 667 LHLEKRMKL--PHAITVMCFGMIFSVTAAYPLVNVASHYHRVPCLPF--DVDITVAKY 722
 QY 733 TVALVMNSFCFIVYAGYIKLYCDLPRGDFEAWDC--AMVRVAVMLIFADGLIXCPV 789
 DB 723 VGSILILNLAVIIMACYASIYLA--OGSH--AMNCNDSRVARRMSLIVFTDFACWAPI 779
 QY 790 AFLSFASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLLENPHRDDLRL 841
 DB 780 AFFSLTAFFGLRLISLDGAKVLTIVLPLNSCANPFLYTLTKOKKCKTI 831

RESULT 12

Q95Y17 PRELIMINARY; PRT: 1280 AA.
 ID 095Y17
 AC 095Y17
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE GLYCOPROTEIN HORMONE RECEPTOR.
 GN AGPBR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 CC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OX NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.:
 RT "cDNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB061861; BAB68208.1;
 KW Receptor.
 SO SEQUENCE 1280 AA; 141700 MW; 0AB0ECC0DD880BA CRC64;

Query Match 20.1%; Score 1017.5; DB 5; Length 1280;
 Best Local Similarity 30.2%; Pred. NO. 5.2e-68;
 Matches 269; Conservative 132; Mismatches 360; Indels 131; Gaps 18;

QY 9 AMLTCAALC-----ASRRAGAPQGPPTACPARCHCOEDGIMLSADCSSELGLASVPD 63
 DB 12 SYVLLLLCLLRVVISRVCVGDGFLCPGLCC--CH--DG-GTHVNCITRRRLDVPAS 64
 QY 64 LDPLTAVYLDLSNNLTLELOPGLFHLRLFEELRLSGNHLSTHPGAFSGLYSLKIIMLON 123
 DB 65 LIGITETIDLSNNITSIILPADAFRIHPRDITLILIGNRLSTIDKNVFRGLRINDITNLKL 124
 QY 124 NOLGSIAPALM--ELPSIQSLRLDANLISLVPERSEGLISRLHMLDNDNALTEIPVA 181
 DB 125 NFFQOVPKRAFRNDLANRLKRLHDSNWIREFVADAFNMLTALHNLHNLHNOISEVPTAA 184
 QY 182 LNNPLALQAMTALNRISIIIPYAFONLSTLYVHLHNNRIGHGHSDEGLNLETLD 241
 DB 185 LHLTSLRLTLHLSHSIPVYPDPAFAENSHLLEILRLHKKITHLSAHAPAGLPNMLLE- 243

QY 242 NNNKLOEPPVALIRLGLRLOELGFHNNNNIKALPEKAMGNPLQTLHFYDNPLOFVGRSAF 301
 DB 244 -----FLGSIISIAH-----TAF 257
 QY 302 QYLPRLHTLSLNGAMDIOEFPDLKGTSTLEITLTRAGIRLLPSGMOQOLPRLRYLESH 361
 DB 258 RNLPLRLNVLILEVNNLSVFPDLTGTSTLEHIGIRCSLRAPANFCNMTGLSTLNLN 317
 QY 362 NOIEELPSLHRCQKLE-----ETGLQHNRIWEIGADTFESQL 397
 DB 318 NLIEGLPSLSCSSLSKVLHLGTLNKLTLSEGGPFGHLDYDLOLENDISTYIPALAFOSL 377
 QY 398 SSLQALDLSWNAIRSIHPEAFSTLSLVKLDLTNQLTTLPLAGLGLMHLKLGNTALS 457
 DB 378 SHLDLTLSTNNITIRELDSQAFAPCTSLQYLDLSNNSFPVLPFAGLOMLKITYDNEQE 437
 QY 458 QAFSKDSEPKRLILEVPYAYOCPPYGMCAEFKASQGWAEEDLHDDDESSKRPGLGLAR 517
 DB 438 DEPPSELPTSTETATATAPYHCCEYIELAEYLS-----LADRPNISETTYWASG 488
 QY 518 QAENHYDDDELQLEMEDS-----KHPVSQCSPTPGPEKP 554
 DB 489 SVDPYNNMTFEDINSWSIDSIFFGSLISIGSPTYLSGNTSRLVPH-NISCRPKRGPEMP 547
 QY 555 CEYFESMGIRLAWAYVLLSVLCNGVLLTVFAGGPARLPVKYVGAIAAGANTLTGIS 614
 DB 548 CMDLFGSMPRLIGWMLFELALIGNAIVFVIYVSHTRKMDVP-REFLGNLAFADFLGY 606
 QY 615 CGLASVVALTFGQSEYKAWETGLGCRATGFLAVLGSEAVLLTLTAAVOCYSVSCV 674
 DB 607 LGFLAGVDTSTLGVPRKRGARQWLSAGCRLAGFLAVFSESESYTLSTLTERFAIKHA 666
 QY 675 RAYGKSPSLGSRVAGVIGCLA--LAGLAALPLASVGEYASPLCLPYAPPEGOPALGF 732
 DB 667 LHLEKRMKL--PHAITVMCFGMIFSVTAAYPLVNVASHYHRVPCLPF--DVDITVAKY 722
 QY 733 TVALVMNSFCFIVYAGYIKLYCDLPRGDFEAWDC--AMVRVAVMLIFADGLIXCPV 789
 DB 723 VGSILILNLAVIIMACYASIYLA--OGSH--AMNCNDSRVARRMSLIVFTDFACWAPI 779
 QY 790 AFLSFASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLLENPHRDDLRL 841
 DB 780 AFFSLTAFFGLRLISLDGAKVLTIVLPLNSCANPFLYTLTKOKKCKTI 831

RESULT 13

Q9NRD6 PRELIMINARY; PRT: 1300 AA.
 ID Q9NRD6
 AC Q9NRD6; Q9YJ03;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HYPOTHETICAL 144.0 KDA PROTEIN (RK GENE PRODUCT).
 GN RK OR BG:DS00180.13 OR CG89350.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Y, AND CN BW SP;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazef R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
 RA Celisner S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [2]

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishu S., Hsu S.Y., Zell K., Hsueh A.J.;
 RT "Characterization of two fly LGR (leucine-rich repeat-containing G
 RT protein-coupled receptor) proteins homologous to vertebrate
 RT glycoprotein hormone receptors: constitutively activation of wild type
 RT fly LGR but not LGR2 in transfected mammalian cells."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DB EMBL: AF274591; AK00808.1; _
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm.1; 1.
 DR Pfam: PF00560; LRR_10.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR SMART: SM00370; LRR; 8.
 DR SMART: SM00369; LRR_typ; 10.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 1050 AA; 117707 MW; 35D71260A8B4BF99 CRC64;
 Query Match 18.6%; Score 941; DB 5; Length 1050;
 Best Local Similarity 27.8%; Pred. No. 2,4e-62;
 Matches 259; Conservative 148; Mismatches 349; Indels 176; Gaps 21;
 OY 104 HIRGO--AFSGVLSLKLTL-----MLNNNGIGPAEALWELP 138
 DB 2 HLPNNYLFSEHFIRSLFSPVLYGFCNGYKKRKTNNPHQNLGNALVSDGCLGHQ 61
 OY 139 SLQSLRLDANLISVPERSEFGLSLRHLMDNALTEIPYRALNMLPALQAMTLALNRI 198
 DB 62 KLRLRLRLEGNL-----FYRIPLNALAGRLTEALNIGSNIL 97
 OY 199 SHIPDYAFQNTLSLVYLHLNHNRIOLGTHSEFEGHANETDLNANKIOEPFVALRTIGR 258
 DB 98 TLINDEDEPRPRLNVLVLLKRNQIKKISAGALKNLTKLVELDNLNLSSEGLSKLSQ 157
 OY 259 LOELGFHNHNKAIPEKAFMGNPLOTIHFDPNPIQFVGRSAFOYLKLTLSLNGAMD 318
 DB 158 LOELSTNRLRLWINDTELPSS--WQMDMRANPLSTISPGAFRGMKSKRLKILSDVRL 215
 OY 319 QEPDLKGTSLLEITLRAGIRLLPSGMCQQLPRLVLELSHNOIEELPSLHRC----- 373
 DB 216 RSFPELEACHALEILKLRAGIOEVANLCLROTPLKSLKLTNSIKRIPNLSSCRDLRL 275
 OY 374 -----QKLEELGLOHNRIMEIGADFEQSLSSLOALDLSMNAIRSIH 414
 DB 276 LDLSNNQIEKIQGKPPSGIKQNLNLLSLNRIKALPQDAFQGIPIQLDLLEGENEISYTH 335
 OY 415 PRAEFTLSHVLKIDLTNQLTTLPLAGLGMLKLGMLKGNLALSOAFSPFKRLILEVP 474
 DB 336 KAFSGFALDELNLGNINFPPLPESGLRALHLKTFENPKLRREPPTPTPRIGLILIS 395
 OY 475 YAYQCC---PYGMCASFKKAS----- 492
 DB 396 YAYHCCAFPLIVAMSSOKKTSQVOEAVLPSPDAEFTTLNNSNMNIWPMHNLKOLGA 455
 OY 493 ---GQMEADLLDDESSKRPRLGLLAROENHYDQDLDE-----LQLEME 535
 DB 456 SMHDPRETA--INNEEQLOTOTGGQIATISYMEYFEENHVSQPAITGYGTGLFSQMSRE 514
 OY 536 DSKPSPVOCSPPTGPFKPCXYLFEWSGIRLAWAVLTLVLCNGLVLLTVFAGPAPPL 595
 DB 515 DQGP--GSVQCLPMPGPFLCADFLDMWTLRGVWVVFLLSLGNGTVFVLLC--SRSKMD 572
 OY 596 PVKFTVGAITAGANTLTGICGLASDALTFQGFSEYSGARWETGLCGRATGFLAVLGSEA 655
 DB 573 VRFVLVCLNLAADDFWGIYLGIIAIVDAATLGEFRMAFIPWQMSVLCQLSGFLAVLSSEL 632

OY 656 SVLLTLAAVQCSVSCVRAYGKPSLSGYRAGVLCGL--ALAGLAALPLASVGEYGA 713
 DB 633 SV--YTLAIVTLERNAVITRAIHLNKRSLSKQAGIYMSVGVFRLIAMLPLVGSYRK 690
 OY 714 SPLCLPYAPPEGGPAPALGFVALVAMNSFCFLVVAAGVYIKLYCDLPGRDGEAVW---DCA 770
 DB 691 FAVOLPPEYTTTG--PASLTVYVLSLFMFINGCAFLTLMGCLKMYAI-RG--SQAWNTDSR 746
 OY 771 MVRHVAALIPADGILYCPVAFLSASHLGLPVPYRPAVKSVLYVLRPLACLPNLYLLEF 830
 DB 747 IAKRMALVLFDFLCWSPFIAPFSITAFGLDLSLEQAKITVVLPLNLSGNPFLYAIM 806
 OY 831 NPFRDRLRLRPR-----AGDSGLAYAAAGELEKSSCDSTOALVAFSDVDL---ILE 881
 DB 807 TKQFKKDCVTLCKHFEESRYVGGGPGRGAVARTKRQ-----DLPPILP 852
 OY 882 ASEAGRPPELTGYGPPSVTLISCOOPGARPLE 913
 DB 853 AAAYAHPPGCR-----CLRMPLSEMPNMHKNME 879
 RESULT 15
 Q9DGC6 PRELIMINARY; PRT; 701 AA.
 AC Q9DGC6;
 DT 01-MAR-2001 (TREMBLrel. 16, created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE GONADOTROPIN RECEPTOR 1.
 GN TGNH-RL.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorphi; Acanthopterygii; Perciformes; Labroidae;
 CC Clariidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.;
 RT "Tilapia gonadotropin receptor I";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB041762; BAB16106.1; --
 DR HSSP: P23945; 1XON.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm.1; 1.
 DR Pfam: PF00560; LRR_5.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 701 AA; 78453 MW; B3D78465CA56410A CRC64;
 Query Match 10.8%; Score 546; DB 13; Length 701;
 Best Local Similarity 26.8%; Pred. No. 1.1e-32;
 Matches 177; Conservative 112; Mismatches 310; Indels 62; Gaps 21;
 OY 226 GTHSF--EGLHNLETLNANKLOEFPV-AIRTLGRLOELGFHNHN--IKAIPEKAFMGNP 282
 DB 51 GYTAFFSNISNAQCLEVKQTQRIHQOGTILSSLOHLMELTISNDLESIGARAFGLPH 110
 OY 283 LOTIHFYDN--PIQFVGRSAFOYLKLTHTLSLNGAMDIOEF--PD-LKGTSLLEITLRAG 339
 DB 111 LTKILISKNAALNIGAFVFSNLPSELSEIITTSKHLSEFIHPDAFRMARLRFLLTSNTG 170
 OY 340 IRLPBGMCQQLPRLVLELSH--QIEELPS--LHRC--QKLEELGLOHNRIMEIGADTF 394
 DB 171 LRIFPFSKIHSTACFLDLQDNSHIRKVPANAFRGICLYCTFAEIRLTRNGIEVASDAF 230
 OY 395 SOLSSLOALDLSWN--AIRSIHPEAFSTHSLVKIDLTNQLTTLPLAGLGMLKIKGN 453
 DB 231 NG--TKMHRFLGNGRQLTHISPAVSSSELVLDVSEFATLSLPDSIIDGLKRLIASEA 289

